



OPEN

DATA DESCRIPTOR

A benchmark dataset for binary segmentation and quantification of dust emissions from unsealed roads

Asanka De Silva¹, Rajitha Ranasinghe¹, Arooran Sounthararajah¹, Hamed Haghighi² & Jayantha Kodikara¹✉

The generation of reference data for machine learning models is challenging for dust emissions due to perpetually dynamic environmental conditions. We generated a new vision dataset with the goal of advancing semantic segmentation to identify and quantify vehicle-induced dust clouds from images. We conducted field experiments on 10 unsealed road segments with different types of road surface materials in varying climatic conditions to capture vehicle-induced road dust. A direct single-lens reflex (DSLR) camera was used to capture the dust clouds generated due to a utility vehicle travelling at different speeds. A research-grade dust monitor was used to measure the dust emissions due to traffic. A total of ~210,000 images were photographed and refined to obtain ~7,000 images. These images were manually annotated to generate masks for dust segmentation. The baseline performance of a truncated sample of ~900 images from the dataset is evaluated for U-Net architecture.

Background

Unsealed roads are the largest component of the road network in many countries, including Australia, New Zealand and South Africa, where unsealed roads amount to approximately 60%, 40%¹ and 75%² of all roads, respectively. Dust emission from these roads is a serious issue, as it has adverse health and environmental impacts^{3,4}. Furthermore, the generated dust cloud reduces road visibility⁵ leading to traffic hazards^{6,7}. Dust is essentially the loss of material from the road surface, which is an indication of the degree of deterioration of an unsealed road⁸. In order to determine the most appropriate maintenance strategy and minimise dust emissions, the dust generated needs to be quantified. Multiple methods exist to measure dust emissions from unsealed roads. The AP-42 dust model developed by the United States Environmental Protection Agency (USEPA) estimates the quantity of particles with diameters less than or equal to 10 μm (PM_{10})⁹; however, this empirical model shows discrepancies with *in-situ* dust measurements¹⁰. There are various field-based methods for physical measurement of dust in terms of mass or volumetric density. These methods vary from vacuum pumps to light-scattering systems^{11–14}. Recently, several research groups have investigated the performance of different machine learning (ML) methods in detecting dust in large-scale dust events observed by satellites utilising different dust spectrum signatures such as empirical thresholds, dust false colour imaging, dust index^{15–17}, deep blue algorithm¹⁸ and K-means clustering¹⁹. However, the applicability of these methods to localised dust events, such as dust emissions from unsealed roads is still questionable. Further, the sensitivity of these methods to spectral bands used during training has not yet been examined. In a recent study, encoder-decoder ML models were used to perform dust segmentation for satellite images of dust storms on Mars²⁰. However, more advanced, recently-developed ML models with more components than encoder-decoder models could be used for more accurate dust segmentation. Another study used smartphones to gather images of dust emissions from unsealed roads and collected visual indicators for further image analysis using different filters; however, only dust classification and dust severity were reported, and the dataset is not publicly available²¹. Researchers have also explored different classification methods for dust detection in satellite data, such as vector machines and random forests²², and found that the algorithms perform better than the statistical techniques²³. Further, a classification study was carried out on dust emissions in a coal preparation plant by localising particle overlapping regions and feature learning through a discriminatory network²⁴. Unlike classification techniques, semantic segmentation has an

¹ARC Industrial Transformation Research Hub (ITRH) — SPARC Hub, Department of Civil Engineering, Monash University, Clayton Campus, Clayton, VIC, 3800, Australia. ²National Research and Development Laboratory, Downer EDI Works, Somerton, VIC, 3061, Australia. ✉e-mail: Jayantha.Kodikara@monash.edu

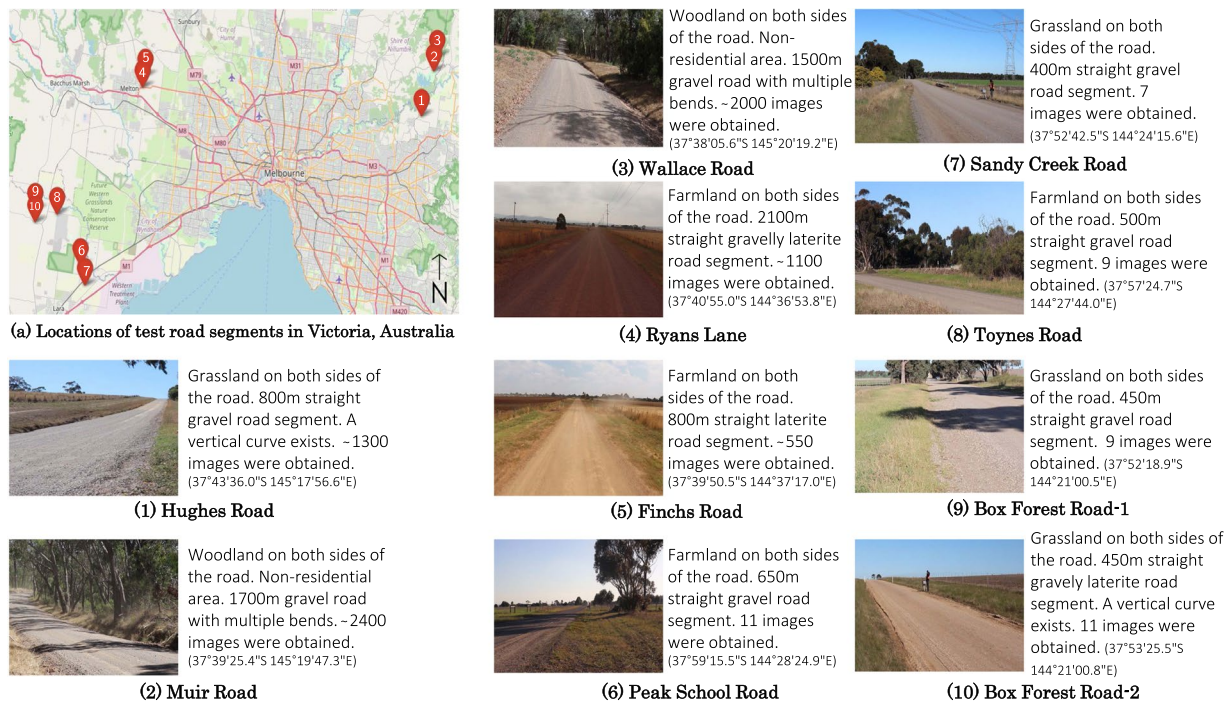


Fig. 1 Details of the test road segments and the nature of the surface of each road segment. Images from Hughes Road, Muir Road, Wallace Road, Ryan's Lane and Finchs Road were used for training and the images from Peak School Road, Sandy Creek Road, Toynes Road, Box Forest Road-1 and Box Forest Road-2 were used to produce maximum dust clouds.

advantage as it could distinguish the boundary of the object and the object area that requires identification, providing semantic segmentation with an upper hand on quantification problems. A reliable dust measurement technique should be able to recognise different sizes of particles. Image processing and analysis cannot do this directly, as most particle sizes are too small to capture in a single pixel. However, if the particles combine together (e.g. in a dust cloud), pixel-wise dust detection of an image becomes possible²⁵. Past studies have used supervised machine learning methods to recognise dust pixels in satellite images. Trained models are then used to determine whether an observed pixel on the satellite image is dust^{26–28}.

The application of supervised machine learning tools for the detection of dust from unsealed roads is still very new. The major limitation of supervised machine learning is the requirement of a large training set of images²⁹ from unsealed roads in which dust pixels are labelled. The manual pixel-wise labelling of images is labour-intensive and very time-consuming³⁰. However, a dataset consisting of raw and labelled images remains essential for use in machine learning techniques to accurately identify and quantify dust. Our dataset offers ~7,000 ground truth images and ~7,000 annotated images in a benchmark set, enabling its use for training deep learning models to recognise the dust pixels on unsealed roads.

Methods

Field experiments were designed to gather images of vehicle-induced road dust with their corresponding dust concentrations. Ten unsealed road segments in Victoria, Australia were selected for the conduct of the experiments.

Test locations. Figure 1 provides details about the roads and a picture of the surface of each road segment.

Test setups. Data were collected in mobile and stationary test set-ups for vehicle speeds ranging from 10 km/h to 90 km/h depending on safety considerations and the local road rules.

In the stationary set-up, the camera and the dust monitor were mounted on the side of the road. Two individual test configurations were applied, depending on the positioning of the camera. The camera either had a longitudinal view of the road or a side view of the road as the test vehicle was driven past it. The two test configurations are shown schematically in Fig. 2a,b. The actual experimental set-ups for the stationary configuration are shown in Fig. 3a,b.

In the mobile set-up, the vehicle was driven along the roads with the camera and the dust monitor attached to the rear, capturing the background of the rear view. The schematic and the actual test set-ups are shown in Figs. 2c, 3c, respectively.

Dust measurement. A research-grade real-time dust monitor, the DustTrak™ DRX Aerosol Monitor 8533 from TSI, was used in this study. It is able to simultaneously measure both mass and size fraction continuously every second using the light-scattering laser photometric method³¹. Dust measurements in terms of particulate

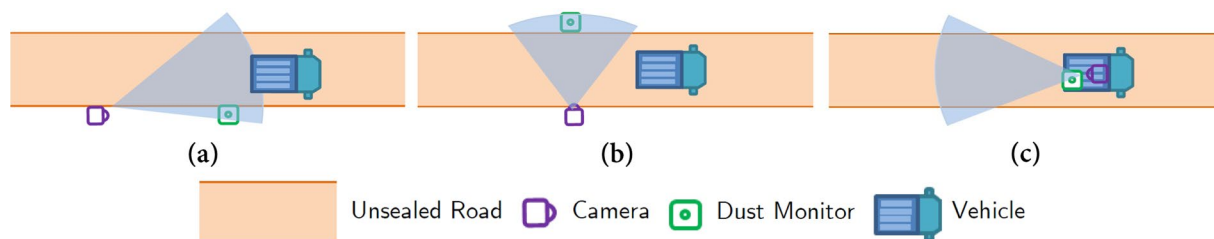


Fig. 2 Schematic diagram of test set-ups. (a) shows stationary set-up with camera capturing longitudinal view of the road, (b) shows stationary set-up with camera capturing side view of the road, and (c) shows mobile set-up with camera attached to vehicle capturing rear view while in motion.

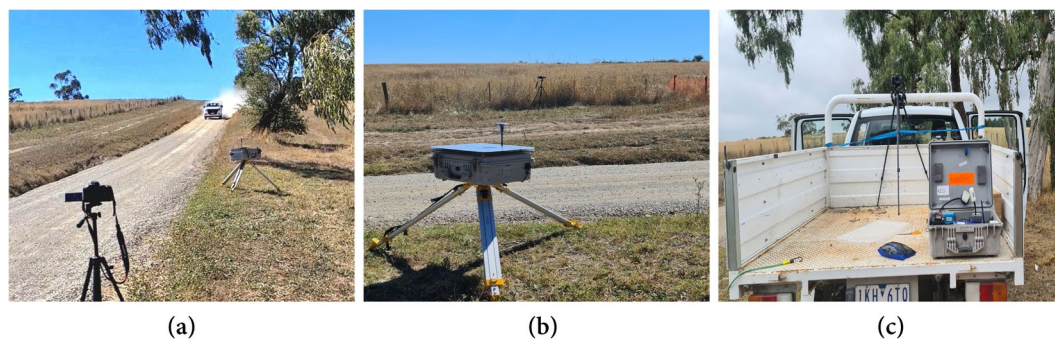


Fig. 3 Real arrangement of instruments at sites. (a) shows stationary set-up with camera capturing longitudinal view of the road, (b) shows stationary set-up with camera capturing side view of the road, and (c) shows mobile set-up with camera attached to vehicle to capture rear view while in motion.

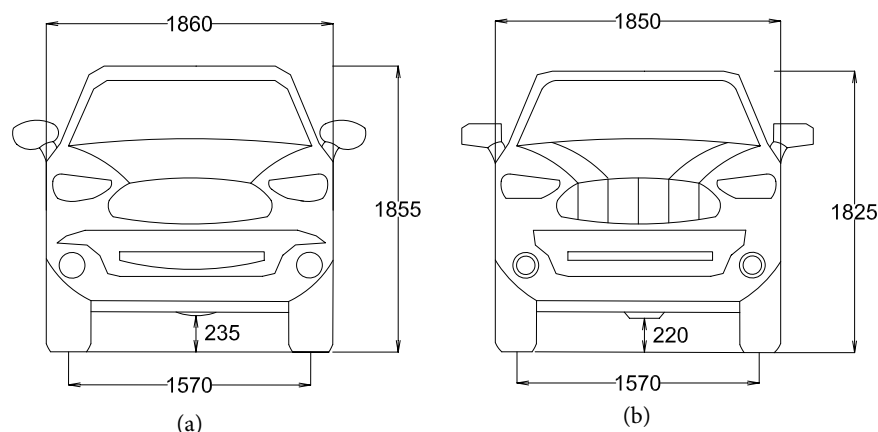


Fig. 4 Cross-sections of the test vehicles used for field experiments. (a) Isuzu D-Max, (b) Nissan Navara ST. All dimensions are in millimeters.

matter (PM_x) of different sizes such as PM_1 , $PM_{2.5}$, PM_4 , PM_{10} and total particulate matter were obtained from the dust monitor for each experiment. (Note: PM_x refers to particulate matter less than or equal to x microns). In future studies, these dust measurements will be used in conjunction with image data to develop a pragmatic tool to quantify vehicle-induced dust from images.

Test vehicle. An Isuzu D-Max or a Nissan Navara ST was used for field experiments, and their cross-sections are provided in Fig. 4^{32–35}. Other characteristics of the test vehicles are provided in Table 1.

Image collection. The images were collected from videos captured by a Canon EOS 200D direct single-lens reflex (DSLR) camera mounted on a tripod at fixed focal length and 25 frames per second (fps). The camera was positioned according to the test set-up. Multiple angles and positions for the camera were selected to capture most of the dust cloud before it dispersed as well as background features. The dust monitor was placed in the wind-dominant direction so that the dust cloud reached the sampling inlet of the dust monitor. Video recording

Property	Isuzu D-Max	Nissan Navara ST
Ground clearance unladen (mm)	235	225
Kerb weight (kg)	1860	1944
Overall length (mm)	5295	5120
Aerodynamic drag coefficient	0.47	0.37

Table 1. Properties of test vehicles.

was started prior to the vehicle beginning to move in order to capture the environment before the dust cloud formed. Therefore, the dataset includes images of the emergence, dissipation and disappearance of the dust cloud, capturing the entire phenomenon. The inclusion of images without dust avoids over-fitting of the dataset.

Image selection. The videos were captured at 25 fps. We selected one image from every 25 frames in a given second and attributed the selected image to that particular second. The image was selected in order to have a unique data point based on the quality of the image by visual inspection. Images with issues such as blur due to motion, etc. were discarded.

Annotation. The task of annotation of an image requires that each pixel of the image be labelled as belonging to dust or non-dust. The annotation process starts with converting the selected RGB image to a gray-scale image, dust is then labelled in white and everything else is labelled in black, generating an image with black and white pixels in different shades. Care was taken to detect very minor to stark differences in dust clouds and accurately annotate them, including transparent regions such as the dust cloud boundary and background. Pixel-wise annotation was performed by two mid-level engineers and verified by a senior engineer and a senior academic supervisor. The annotated images were also reviewed by a senior engineer, who is an expert in dust monitoring in unsealed road environments, representing road contractors. It is apparent that there is no definite boundary for dust clouds; however, in the annotation process, we demarcated the boundary visible in the image. The overall annotation process took 1500 + hours.

Image post-processing. To make manually-annotated images more machine learning-friendly, the images need to be binarized based on pixel intensities. This was done using Otsu's thresholding method³⁰ to generate a new binary image known as a segmentation mask.

Let $\mathcal{X} = (\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i)_{i=1}^n$ be a labelled set with n number of samples, where each sample $(\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i)$ consists of an image $\mathbf{X}_i \in \mathbb{R}^{C \times H \times W}$, the corresponding annotated image $\mathbf{Y}_i \in \{0, 1\}^{C \times H \times W}$ and its thresholded image $\mathbf{Z}_i \in \{0, 1\}^{H \times W}$. Each \mathbf{X}_i is resized so that the height (H) = 1024 pixels, the width (W) = 1024 pixels and the class (C) = 3 for RGB images. Pixels with 0 and 1 represent the non-dust and dust pixels, respectively. Then, $\mathbf{Z}_i = \text{Otsu}(\mathbf{Y}_i)$ where $\text{Otsu}()$ is the Otsu's thresholding function.

Data Records

The URDE dataset is available from Figshare (<https://doi.org/10.6084/m9.figshare.20459784>)³⁶. In the repository, there are three folders, *RandomDataset_897*, *SequentialDataset_7k* and *Dust Readings*. *SequentialDataset_7k* folder has all the images from the experiments and a Google search. *RandomDataset_897* folder has two folders containing 800 images in the *Training* folder and 97 images in the *Validation* folder. The original dataset is *SequentialDataset_7k*. *RandomDataset_897* is a sub-sample of *SequentialDataset_7k*, selected based on the similarity of consecutive images. To minimise undesirable over-fitting phenomena in machine learning and reduce bias, visually similar consecutive images were removed. *RandomDataset_897* is a representative of the larger dataset *SequentialDataset_7k* and it is sufficiently large to segment road dust with high accuracy and produce optimal results. The *Dust readings* folder has dust images corresponding to the maximum dust cloud and a spreadsheet containing the maximum dust concentrations corresponding to that dust cloud. We currently use the dust concentration data to develop a practical dust prediction model for use in the maintenance of unsealed roads. For each image in the folder, an annotated image and a segmentation mask exist in respective sub-folders. *RandomDataset_897* was included so that any researcher could reproduce our results or reduce unnecessarily longer training time upon checking for the trainability of new ML models. *SequentialDataset_7k* is included for future works, in particular, where sequential images are advantageous. Figure 5 illustrates the file structure of the repository.

Technical Validation

The original images, annotated images and produced secondary images were validated for quality and trainability using the U-net architecture. U-net was selected as it performs moderately well for binary segmentation tasks across multidisciplinary datasets, and it is the root architecture for many modern architectures such as DenseUNet³⁷. In Fig. 6, we present qualitative results for our dataset with baseline variables of the architecture and hyper-parameters. We conducted a comprehensive analysis of the dataset for multiple state-of-the-art ML algorithms, and the results showed that the accuracy of segmentation of dust by different ML models increases from the pioneering vanilla Unet to more advanced architectures such as DeepLabV3.

Segmentation performance evaluation. To characterise the results of our experiments, we chose the Dice Similarity Coefficient (DSC)³⁸ (Eq. 1) and the Loss. The two parameters are plotted against the number of

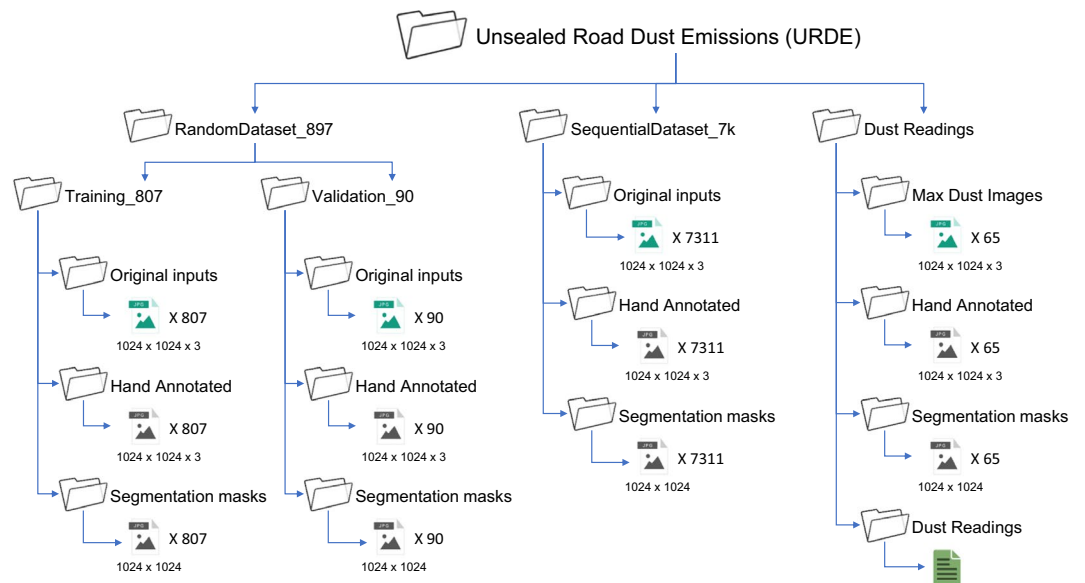


Fig. 5 Data repository structure.

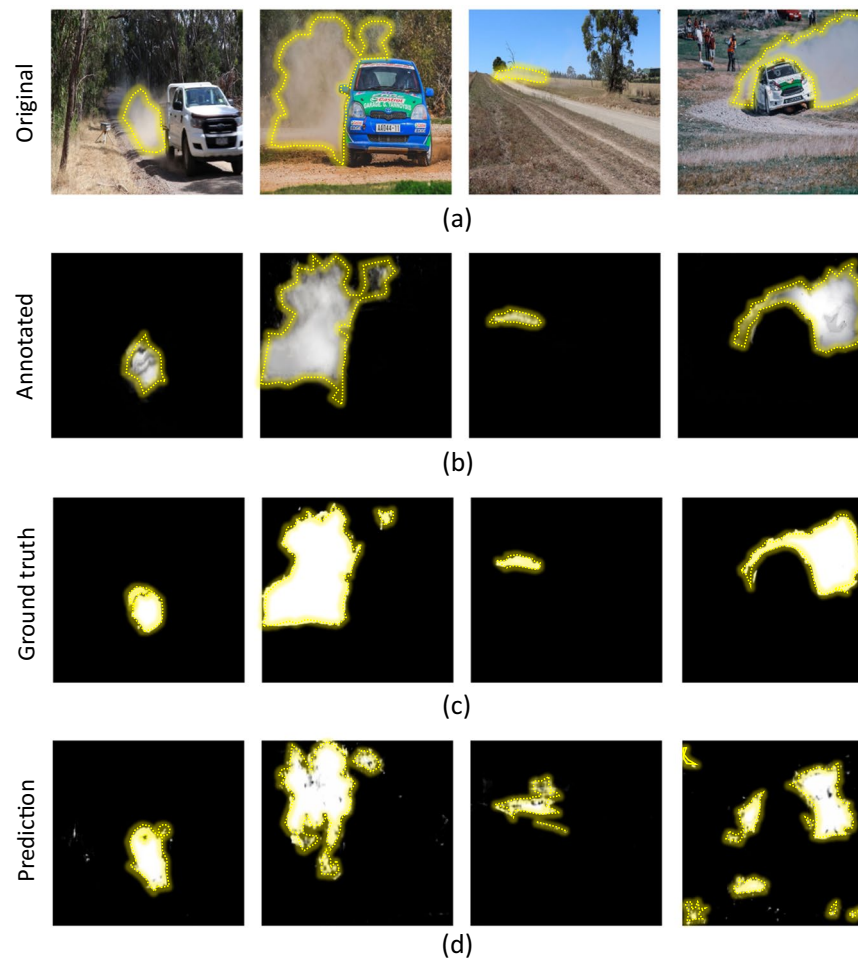


Fig. 6 Visualisation of segmentation. (a) Original images from the dataset, (b) Manually annotated images, (c) Images segmented using Otsu's thresholding, (d) Dust cloud predicted by U-Net. Dust cloud boundary is demarcated by yellow dotted lines in all images.

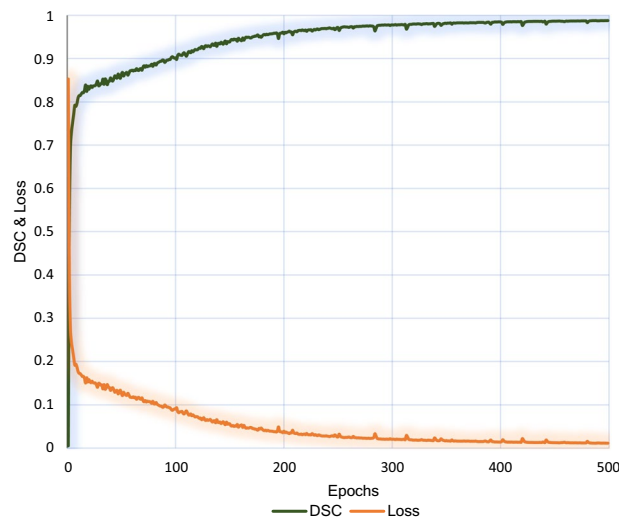


Fig. 7 DSC and Loss.

Segmentation model	Validation DSC	No. of Epochs	Learning rate	Batch size	Model input resolution
U-net	0.906	500	$5 \times e^{-4}$	4	256×256

Table 2. Quantitative results and Hyper-parameters.

	GPU	Video-RAM	System-RAM
Platform	Nvidia 3080 mobile	16GB	32GB
Usage	Fully	5.2GB	11.5GB

Table 3. Training Platform.

epochs for the training as shown in Fig. 7. Table 2 shows the quantitative results and hyper-parameters from evaluation. The training platform used for evaluation is described in Table 3.

$$DSC(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (1)$$

where, A and B are sets of voxels of ground truth and the segmented images, respectively. TP, FP and FN are the total number of true positive voxels, false positive voxels and the false negative voxels, respectively.

For training, the batch size and number of epochs were selected to be 4 and 500, respectively. The batch size of 4 means that four images are processed before the internal model parameters are updated. By making the number of epochs 500, the model does 500 complete iterations through the training dataset. The input layer of the ML model accepts an RGB image vector with 256×256 resolution. We selected 256×256 resolution conservatively to get the best performance without having a bulkier ML model because high input resolutions exponentially increase the model parameters and subsequently the model size without any added benefit to precision.

Precision and repeatability of data. In the annotation process, randomly sampled images from the dataset were annotated independently by two annotators. The annotated images were reviewed independently by a senior engineer and a senior academic supervisor. Similar image sets annotated by the two annotators were checked for trainability using U-Net architecture, and the DSC was evaluated for comparison to ensure all secondary data are repeatable.

Limitations. Several limitations of the proposed method were identified as follows:

1. The dataset contains images where the background or the road in the image does not complement the appearance of the dust cloud due to its visual similarity. This results in a “camouflage effect” where the model is unable to distinguish between the dust cloud and the background or the road. Therefore, a part of the background or the road may be misrecognised as dust.
2. A majority of the images in *RandomDataset_897* folder includes ~800 images collected from experiments conducted in Victoria, Australia. The 897 images include 85 copyright-free images from Google, 189 images

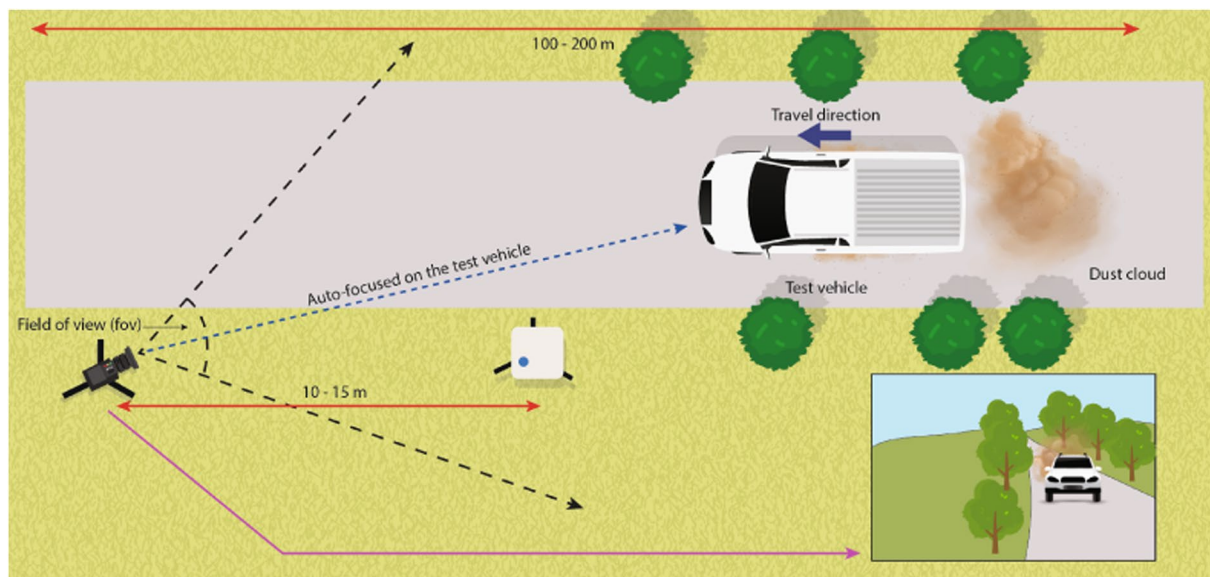


Fig. 8 Boundary conditions of the field experiments.

from Hughes road, 219 images from Muir road, 209 images from Wallace road, 112 images from Ryans lane and 83 images from Finch road, respectively. The 85 copyright-free images were added to increase variability. However, due to the disproportion between the number of images from experiments and the images from Google, the model currently fits best for unsealed roads in Australia.

3. In the binarizing step, the regions where the dust cloud was transparent are identified as dust pixels leading to over-estimation of dust cloud from the image.
4. To ensure the precision and repeatability of the data collected in field experiments, we established experimental parameters and boundary conditions shown in Fig. 8. The Canon EOS 200D DSLR camera was set to auto-focus the oncoming vehicle (Isuzu D-Max or a Nissan Navara ST). The camera position and field of view were selected so that 100 to 200 meters of road length is visible in the video. The dust monitoring device (DustTrak™ DRX Aerosol Monitor 8533 from TSI) was positioned 10 to 15 meters from the camera. The resolution of the video and image was set to 1920×1080 pixels. The traffic-induced dust emission in unsealed roads is affected by many factors, which include wind speed and direction, condition of the surface material, vehicle tire type and its condition, etc.; however, similar data could be reproduced if the field experiments are conducted with the experimental parameters and boundary conditions shown in Fig. 8.

Limitations 1 and 2 may be resolved by expanding the URDE dataset, which is publicly available at <https://doi.org/10.6084/m9.figshare.20459784>. Limitation 3 may be resolved using a different thresholding technique.

Usage Notes

The dataset includes original images, manually-annotated images and the corresponding binary images generated by Otsu's thresholding method. The binary images generated from Otsu's method can be obtained using the Python-based tool included at https://github.com/RajithaRanasinghe/Automatic_Thresholding. The set of original images together with their corresponding binary images may be used in conjunction with machine learning architectures to train the dataset and predict dust in an image. The dust measurements from the dust monitor may be used to correlate the dust cloud with actual dust readings. The process may be extended to videos and real-time dust prediction applications.

Code availability

All the Python scripts used to generate the secondary data (binary images by Otsu's thresholding) are provided at https://github.com/RajithaRanasinghe/Automatic_Thresholding.

Received: 10 August 2022; Accepted: 19 December 2022;

Published online: 05 January 2023

References

1. Austroads. Asset management of unsealed roads: Literature review, LGA survey and workshop (2000–2002). Report, Austroads (2006).
2. Paige-Green, P. Local government note: new perspectives of unsealed roads in South Africa. In *Low Volume Roads Workshop 2007*. Keynote Address Presented at the REAAA (NZ) Low Volume Roads Workshop held in Nelson, New Zealand, 18–20 July 2007 (2007).
3. Cooper, W. R. Cause and prevention of dust from automobiles. *Nature* **72**, 485–490, <https://doi.org/10.1038/072485c0> (1905).
4. Pope, C. A. & Dockery, D. W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manag. Assoc.* **56**, 709–742, <https://doi.org/10.1080/10473289.2006.10464485> (2006).

5. Baddock, M. C. *et al.* A visibility and total suspended dust relationship. *Atmos. Environ.* (1994) **89**, 329–336, <https://doi.org/10.1016/j.atmosenv.2014.02.038> (2014).
6. Ashley, W. S., Strader, S., Dziubla, D. C. & Haberlie, A. Driving blind: Weather-related vision hazards and fatal motor vehicle crashes. *Bull. Amer. Meteor. Soc.* **96**, 755–778, <https://doi.org/10.1175/BAMS-D-14-00026.1> (2015).
7. Greening, T. *Quantifying the Impacts of Vehicle-Generated Dust: a comprehensive approach*. Quantifying the Impacts of Vehicle-Generated Dust <https://elibrary.worldbank.org/doi/abs/10.1596/27891> (World Bank, Washington, DC, 2011).
8. Pardeshi, V., Nimbalkar, S. & Khabbaz, H. Theoretical and experimental assessment of gravel loss on unsealed roads in Australia. In Kanwar, V. S. & Shukla, S. K. (eds.) *Sustainable Civil Engineering Practices*, 21–29, https://doi.org/10.1007/978-981-15-3677-9_3 (Springer Singapore, Singapore, 2020).
9. USEPA. AP-42, *Compilation of Air Pollutant Emission Factors*, 137–140 (Hoboken, NJ, USA: John Wiley & Sons, Inc, Hoboken, NJ, USA, 2016).
10. Gillies, J. A., Etyemezian, V., Kuhns, H., Nikolic, D. & Gillette, D. A. Effect of vehicle characteristics on unpaved road dust emissions. *Atmos. Environ.* **39**, 2341–2347, <https://doi.org/10.1016/j.atmosenv.2004.05.064> (2005).
11. Merkus, H. G. *Particle Size, Size Distributions and Shape*, 13–42 (Springer Netherlands, Dordrecht, 2009).
12. Hetem, I. G. & de Fatima Andrade, M. Characterization of fine particulate matter emitted from the resuspension of road and pavement dust in the metropolitan area of São Paulo, Brazil. *Atmosphere*, **7**, 31, <https://doi.org/10.3390/atmos7030031> (2016).
13. Amato, F. *Non-Exhaust Emissions: An Urban Air Quality Problem for Public Health Impact and Mitigation Measures* (Saint Louis: Elsevier Science & Technology, Saint Louis, 2018).
14. Lundberg, J., Blomqvist, G., Gustafsson, M., Janhäll, S. & Järnliskog, I. Wet dust sampler—a sampling method for road dust quantification and analyses. *Wat. Air and Soil Poll.* **230**, 1–21, <https://doi.org/10.1007/s11270-019-4226-6> (2019).
15. Ackerman, S. A. Remote sensing aerosols using satellite infrared observations. *J. Geophys. Res.* **102**, 17069–17079, <https://doi.org/10.1029/96JD03066> (1997).
16. Ashpole, I. & Washington, R. An automated dust detection using seviri: A multiyear climatology of summertime dustiness in the central and western Sahara. *J. Geophys. Res.* **117**, D8, <https://doi.org/10.1029/2011JD016845> (2012).
17. Taylor, L., Mackie, S. & Watson, M. Investigating the use of the Saharan dust index as a tool for the detection of volcanic ash in seviri imagery. *Journal of Volcanology and Geothermal Research* **304**, 126–141, <https://doi.org/10.1016/j.jvolgeores.2015.08.014> (2015).
18. Hsu, N. C., Si-Chee, T., King, M. D. & Herman, J. R. Deep blue retrievals of Asian aerosol properties during ACE-Asia. *IEEE Trans. Geosci. Remote Sens.* **44**, 3180–3195, <https://doi.org/10.1109/TGRS.2006.879540> (2006).
19. Strandgren, J., Bugliaro, L., Sehnke, F. & Schröder, L. Cirrus cloud retrieval with MSG/SEVIRI using artificial neural networks. *Atmospheric Measurement Techniques* **10**, 3547–3573, <https://doi.org/10.5194/amt-10-3547-2017> (2017).
20. Ogohara, K. & Gichu, R. Automated segmentation of textured dust storms on Mars remote sensing images using an encoder-decoder type convolutional neural network. *Comput. Geosci.* **160**, 105043, <https://doi.org/10.1016/j.cageo.2022.105043> (2022).
21. Allan, M., Henning, T. F. P. & Andrews, M. A Pragmatic Approach for Dust Monitoring on Unsealed Roads, 439–444. Transportation Research Circular E-C248. <https://www.trb.org/Publications/Blurbs/179567.aspx> (12th International Conference on Low-Volume Roads, 2019).
22. Cai, C., Lee, J., Shi, Y. R., Zervas, C. & Guo, P. Dust detection in satellite data using convolutional neural networks. In *Technical Report HPCF-2019-15* (2019).
23. Kolios, S. & Hatzianastassiou, N. Quantitative aerosol optical depth detection during dust outbreaks from Meteosat imagery using an artificial neural network model. *Remote Sensing* **11**, <https://doi.org/10.3390/rs11091022> (2019).
24. Wang, Z. *et al.* A VGGNet-like approach for classifying and segmenting coal dust particles with overlapping regions. *Computers in Industry* **132**, 103506, <https://doi.org/10.1016/j.compind.2021.103506> (2021).
25. Middleton, W. E. K. *Vision through the atmosphere* (Toronto: University of Toronto Press, Toronto, 1952).
26. Kolios, S. & Hatzianastassiou, N. Quantitative aerosol optical depth detection during dust outbreaks from Meteosat imagery using an artificial neural network model. *Remote Sens. (Basel, Switzerland)*, **11**, 1022, <https://doi.org/10.3390/rs11091022> (2019).
27. Cai, C., Lee, J., Shi, Y. R., Zervas, C. & Guo, P. Dust Detection in Satellite Data using Convolutional Neural Networks, Technical Report HPCF-2019-15 (UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2019).
28. Shi, P. *et al.* A hybrid algorithm for mineral dust detection using satellite data. In *2019 15th International Conference on eScience (eScience)*, 39–46, <https://doi.org/10.1109/eScience.2019.00012> (2019).
29. Yu, M. *et al.* Image segmentation for dust detection using semi-supervised machine learning. In *2020 IEEE International Conference on Big Data (Big Data)*, 1745–1754, <https://doi.org/10.1109/BigData50022.2020.9378198> (2020).
30. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66, <https://doi.org/10.1109/TSMC.1979.4310076> (1979).
31. Dusttrak™ drx aerosol monitor 8533. <https://tsi.com/products/aerosol-and-dust-monitors/aerosol-and-dust-monitors/>, Accessed: 02-08-2022.
32. Isuzu d-max technical specifications. https://s3-ap-southeast-2.amazonaws.com/imotor-cms/files/cms/Isuzu_D-MAX_specification_sheet.pdf. Accessed: 2022-11-26.
33. Sevilla-Mendoza, A. The new d-max: a ride for all seasons and reasons. <https://motioncars.inquirer.net/20825/the-new-d-max-a-ride-for-all-seasons-and-reasons> (2013).
34. Nissan my21 navara dual cab specifications. https://www-asia.nissan-cdn.net/content/dam/Nissan/AU/Files/Brochures/Models/Specsheet/MY21_Nissan_Navara_Spec_Sheet.pdf (2021).
35. Fallah, A. 2010 Nissan Navara ST-X Dual Cab. <https://www.drive.com.au/news/2010-nissan-navara-st-x-dual-cab-utility-upgraded/> (2010).
36. De Silva, A. & Ranasinghe, R. Unsealed Roads Dust Emissions (URDE) *figshare* <https://doi.org/10.6084/m9.figshare.20459784> (2022).
37. Cao, Y., Liu, S., Peng, Y. & Li, J. DenseNet: densely connected UNet for electron microscopy image segmentation. *IET Image Processing* **14**, 2682–2689, <https://doi.org/10.1049/iet-ipr.2019.1527> (2020).
38. Peiris, H., Chen, Z., Egan, G. & Harandi, M. Duo-segnet: adversarial dual-views for semi-supervised medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 428–438, <https://doi.org/10.48550/arXiv.2108.11154> (Springer, 2021).

Acknowledgements

The first author received a Monash Research Scholarship and a Faculty of Engineering International Postgraduate Research Scholarship to undertake this research project. This research work is also part of a research project (Project No IH18.12.1) sponsored by the Smart Pavements Australia Research Collaboration (SPARC) Hub (<https://sparchub.org.au>) at the Department of Civil Engineering, Monash University, funded by the Australian Research Council (ARC) Industrial Transformation Research Hub (ITRH) Scheme (Project ID: IH180100010). The authors gratefully acknowledge the financial and in-kind support of Monash University, the SPARC Hub, and Downer EDI Works Pty Ltd.

Author contributions

A.D., R.R., A.S. and J.K. devised the scientific direction for the dust images dataset. A.D. and R.R. conducted the experiments, analysed the data, produced the secondary dataset and wrote the complete manuscript. All authors reviewed and provided comments on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023