Determination of size of urban particulates from occluded scattering patterns using deep learning and data augmentation

James A. Grant-Jacob[1]*, Matthew Praeger[1], Matthew Loxham[2,3], Robert W. Eason[1] and Ben Mills[1]

1 Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK;
2 Academic Unit of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, SO17 1BJ, UK;
3 NIHR Southampton Biomedical Research Centre, University Hospital Southampton, SO16 6YD, UK;

E-mail: Correspondence: J.A.Grant-Jacob@soton.ac.uk; Tel.: +44-238-059-6975

Abstract

Deep learning has shown recent key breakthroughs in enabling particulate identification directly from scattering patterns. However, moving such a detector from a laboratory to a real-world environment means developing techniques for improving the neural network robustness. Here, a methodology for training data augmentation is proposed that is shown to ensure neural network accuracy, despite occlusion of the scattering pattern by simulated particulates deposited on the detector's imaging sensor surface. The augmentation approach was shown to increase the accuracy of the network when identifying the geometric Y-dimension of the particulates by ~ 62% when 1000 occlusions of size ~ 5 pixels were present on the scattering pattern. This capability demonstrates the potential of data augmentation for increasing accuracy and longevity of a particulate detector operating in a real-world environment.

Keywords: Deep learning, Sensing, Optics, Particulate matter, Pollution.

Introduction

One of the current global health challenges is particulate matter (PM) pollution, since it is associated with sicknesses that include respiratory disorders, cardiovascular diseases and cancer [1–6]. The current World Health Organisation (WHO) guidelines for PM concentrations over a 24-hour period are 25 μg m^{-3} for PM2.5 (< 2.5 μm), and 50 μg m^{-3} for PM10 (< 10 μm particulates), with annual guidelines of 10 μg m^{-3} and 20 μg m^{-3}, respectively [7]. A dataset complied in 2016 by the WHO for annual mean concentrations of PM2.5 and PM10 across the globe, shows that particulate levels can vary from city to city, with 100 μg m^{-3} of PM2.5 and 187 μg m^{-3} of PM10 recorded in Jaipur, India (2012 data), 29 μg m^{-3} of PM2.5 and 64 μg m^{-3} of PM10 recorded in Santiago, Chile (2014 data), and 15 μg m^{-3} of PM2.5 and 21 μg m^{-3} of PM10 recorded in Southampton, UK (2013) [8]. Critically, however, there is no evidence that there is a "safe" lower concentration of PM below which effects do not occur and, indeed, good evidence to the contrary [9]. Mean concentrations mask significant spatiotemporal variation due to proximity to local sources, variation in local and background sources, the effects of meteorological factors such as wind direction, rain and sunlight, and seasonality. There is, therefore, a clear need for PM concentrations in areas of the UK and elsewhere to be monitored. The effect of such pollution on individuals can vary, with some people being more vulnerable to certain types of pollution than others [10]. The size, type and number of particulates can also lead to different adverse health effects. For

example, individual and aggregated particulates larger than PM10, such as pollen grains (some species can have grains as large as ~ 150 μm in geometrical size [11]), can lead to allergic rhinitis [12], and PM10 can cause respiratory problems [13,14]. PM2.5 particulates can travel deep into the lungs, being associated with incidence and exacerbations of respiratory diseases such as chronic obstructive pulmonary disease (COPD) and asthma [15–17], as well as lung cancer [18], while also being shown to exert effects on the cardiovascular system including myocardial infarction and stroke [19]. Ultrafine particulates (< 0.1 μm), a subset of PM2.5, which generally comprise a relatively small proportion of the airborne mass of fine PM, but the majority by number concentration, is thought to be able to translocate from the lung into the circulation, and indeed, ultrafine particles from ambient sources have been observed in both brain tissue and cardiac tissue [20,21], and also appear able to cross the placenta [22]. Importantly, as cohort studies expand in size, and as methods to determine individual exposures improve, associations with further diseases affecting additional body systems are being uncovered [3]. Although these associations are not currently established as being causal relationships, and mechanisms are not fully understood, it is becoming ever more likely that the cardiovascular and respiratory systems are not the sole sites of the deleterious effects of airborne PM [23]. Since particle size and morphology may play an important role in governing interactions with the human body, being able to characterise such airborne particulates is important in determining its prevalence and potential health effects, while also giving possible clues as to its source [24]. Conventional methods for monitoring the size and chemical composition of particulates involve using large filter traps that collect the particulates onto a filter substrate, often with an upstream size selective inlet of impaction surface to allow size fractionation. These filters then need to be post-processed, usually with lab-based scientific apparatus. Collection of particulates in this way clearly reduces the temporal resolution of readouts, and may be labour intensive. A simple optical detector that can determine the size of particulates, whilst also being robust in the environment in which it is sensing, is therefore highly desirable.

The light scattered from a single particulate encodes information regarding its size and shape, as well as chemical constituents and refractive index [25–28]. As shown in the images in the experimental methods section (see figure 1), individual urban particulates are highly variable in shape and size, and consequently the scattering patterns generally also vary significantly. Extraction of particulate parameters via theoretical methods, such as Mie theory [29] can be difficult, due to the complexity of such particulates, and hence alternative methods need to be investigated. In recent years, convolutional neural networks [30–32] have shown notable ability for the classification of objects in images [33–37] and, importantly for this work, they have shown the capability to identify spherical and non-spherical objects directly from their scattering patterns [38–40]. Since a neural network's accuracy is dependent on the data used in training the neural network, methods for improving accuracy include data augmentation, which has been shown to improve neural network accuracy in classifying image and acoustic data [41–43].

Here, we demonstrate the ability to determine the geometric size of the urban particulates from their scattering patterns captured using a complementary metal-oxide-semiconductor (CMOS) camera sensor, via the use of convolutional neural networks. In addition, in optical particle sensing, there is an inherent risk that particulates may deposit on the sensor or on optical surfaces, thereby reducing their ability to collect the scattered light. This is particularly problematic in optical dust detectors, which often rely on just a single photodiode. We demonstrate here that the combination of a camera sensor and a convolutional neural network offers robustness against partial sensor occlusion - potentially improving longevity and accuracy in the event of sensor surface contamination. Current methods for robustness include 'pixel attacking' and 'adversarial attacking' techniques [44,45] that been used in various forms to test and fool neural networks in order to find their weakness, and thus enable improvements in network training. For practical purposes, we use zero intensity pixels of various sizes on the scattering pattern images to simulate particulates deposited on the camera sensor that occlude light from being incident on the sensor's pixels. We improve the accuracy for determining particulate sizes from occluded scattering patterns by augmenting the neural network training data via inclusion of

occluded scattering pattern images, hence demonstrating the potential for improving the detector's robustness when applied in a real-world environment.

Experimental methods

Sample fabrication: Urban particulate matter (NIST 1648a urban particulate matter, collected from St. Louis, MO in 1977-1978) was deposited onto a soda-lime glass slide (25 mm by 75 mm, 1 mm thick). A total of 112 particles were imaged using a Zeiss Evo scanning electron microscope (SEM) and their geometrical sizes, ranging from 3 µm to 45 µm, were determined (maximum X-dimension and maximum Y-dimension in microns, as shown in figure 1(b)). The sizes were chosen for ease of imaging and alignment, with the smallest size limited by the need for sufficient signal to be obtained from the scattering in our experimental setup.

Experimental setup: Figure 1 shows an illustration of the experimental setup, showing light from lasers operating at 450 nm, 520 nm and 635 nm wavelengths, with < 1 mW average powers, being focussed using an Olympus SLMPLN 50× objective (NA = 0.35, WD = 18 mm) onto the surface of the urban particulate-covered glass slide, with a focus diameter of ~ 30 µm. Three lasers were used rather than one since the light scattered from an object is wavelength-dependent, and so the more wavelengths of light used, the more information that is available to the neural networks to allow for more accurate training and prediction, as demonstrated in our previous work [46]. When the lasers illuminated the particulates, the resulting scattered light was collected by a CMOS colour camera sensor (Thorlabs DCC1645C, 1280 × 1024 pixels, 3.60 µm square pixels, 5 ms integration time), that was placed 3 mm further along the laser axis from the illuminated particulates. The sample was mounted onto a 3-axis stage (25 mm maximum travel and 10 µm resolution) to allow alignment of the particulates within the laser beam. To produce in-situ images of the particulates, allowing identification and cross-referencing with the SEM data, they were illuminated with the aid of a white light source (50 W white halogen lamp), and a beam splitter was used to direct the back-reflected light onto a CMOS camera sensor (Thorlabs, DCC3260C, 1936 × 1216 pixel, 5.86 µm square pixels, 5 ms integration time). In total, 113 scattering patterns were recorded, as this included images of scattering patterns from the 112 particulates and 1 image when no particulates were present.

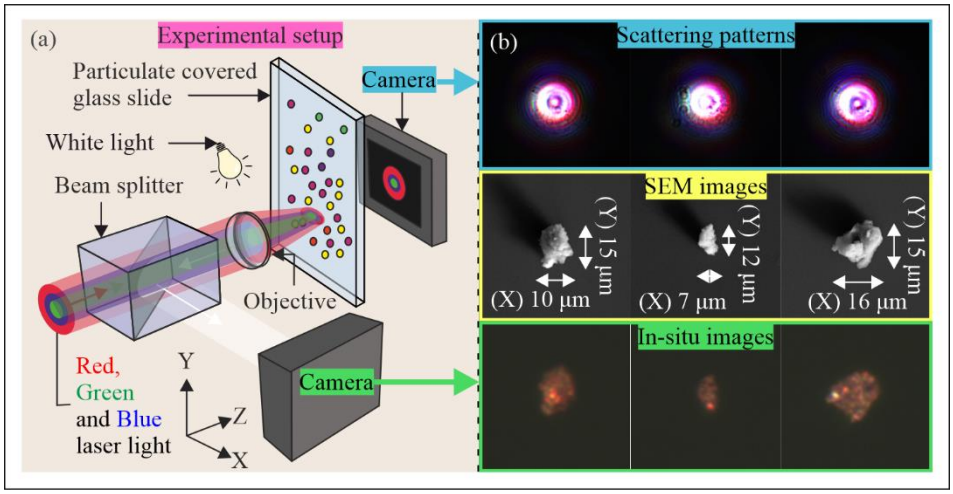

Figure 1. (a) Illustration of experimental setup for using a camera to collect scattering patterns from urban particulates deposited onto a microscope slide, when the particulates were illuminated with red, green and blue laser light. The particulates were also imaged for reference and sample alignment. (b) Scattering patterns and, corresponding SEM images and optical images for three different particulates (upper, middle and lower rows, respectively). The SEM images include labels of the maximum X-dimension and maximum Y-dimension of the particulates.

Neural networks: In total, 3 neural networks were trained using, 1) X-dimension data, 2) Y-dimension data, 3) augmented Y-dimension data. For each network, an 18-layer regression convolutional neural network that included 4 convolutional layers (3×8, 3×16, 3×32, 3×32), each with a stride of [1 1], each followed by a batch normalisation layer and a ReLu (rectified linear activation function) layer, was used. The dataset of training and validation scattering pattern images (total of 92 images) was split in a ratio of 90:10, and the remaining 21 particulate scattering patterns that were not present in the training data and included sizes that were not present in the training dataset (5-25 µm), were used for testing. The input images were cropped to 1024 × 1024 pixels and resized to 256 × 256 pixels, prior to being fed into the neural network. The output consisted of a single number likely between 0 and 45 microns (since this was the range of the values in the training and test data). A learning rate of 0.0002 and a dropout of 0.2 was used, along with a minibatch size of 30, and validation frequency of 5. All networks were trained for 30 epochs using an NVIDIA RTX 2080 graphics processing unit (GPU), taking < 5 minutes each. Initially, the training data and the test data used were unedited, then to test the ability of the neural network to determine particulate sizes for scattering patterns with occlusions on, single zero intensity (black) pixels numbering from 1 to 1000 were randomly added to the test images (5 times each particulate size, for averaging purposes) and the outputs determined. Also, 5 and 10-pixel diameter sized black circles in numbers totalling 1, 10, 100 and 1000 were randomly added to each scattering pattern in the test data and the outputs were again determined. Finally, to improve the ability to determine the particulate sizes from occluded scattering patterns, a neural network was trained using a dataset that was augmented with randomly positioned occlusions. This was achieved by randomly adding 1, 5 and 10-pixel diameter black circles to the training data, in numbers totalling 1, 10, 100 and 1000. Again, the neural network was tested using the same scattering patterns as those tested with the neural network trained without augmented data, for ease of comparison.

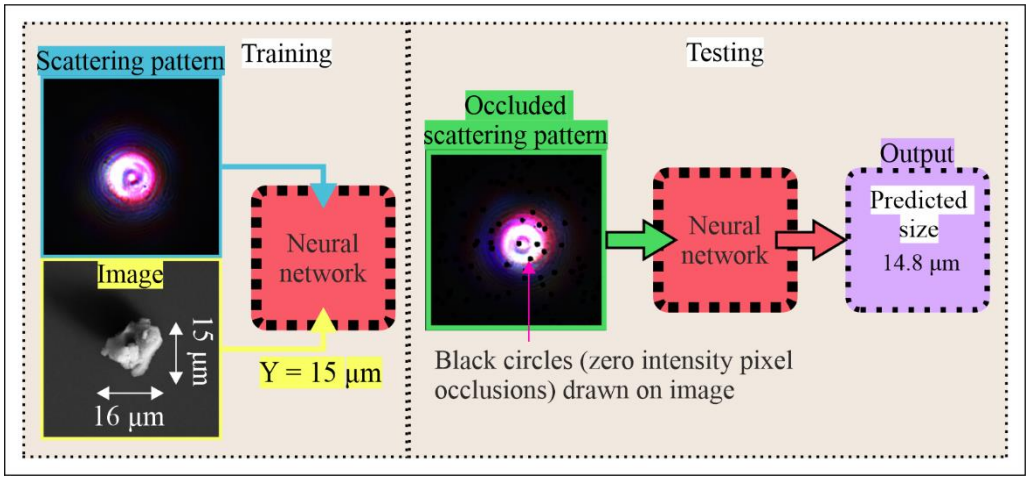

Figure 2. (a) Illustration of the concept of the training of a neural network using a scattering pattern paired with a dimension (e.g. Y = 15 µm), and then testing the network on a scattering pattern. Here, the test scattering pattern includes black circles (zero intensity pixel occlusions) to simulate particulates occluding the sensor.

Results and discussion

The capability of neural networks to determine the size of urban particulates from images of their scattering patterns is shown in figure 3. Evident from the plot is that both the (a) X-dimension and (b) Y-dimension were able to be determined, with an RMSE of 4.00 µm and 3.27 µm, respectively.

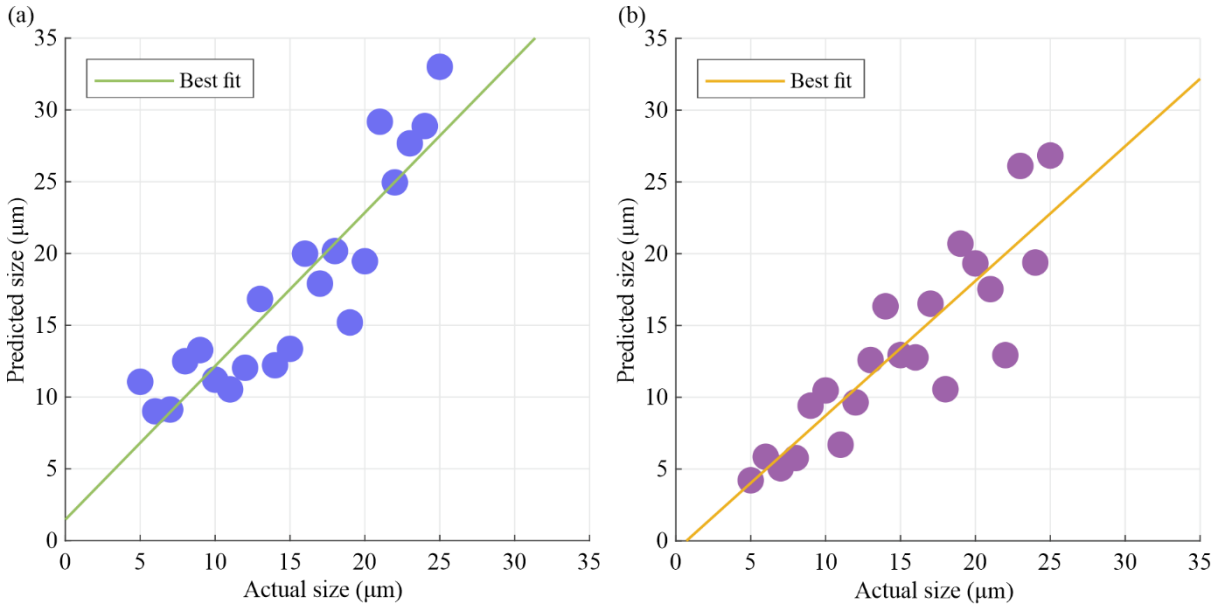

Figure 3. Capability of the trained neural networks to determine the size of the urban particulates for (a) X-dimension and (b) Y-dimension.

To test the robustness of the network against variations in the scattering pattern via occlusion of pixels, figure 4(a) demonstrates the capability of the neural network to predict Y-dimension of particulates, when the scattering pattern images have 1, 10, 100 and 1000 random single pixels set to zero intensity. The figure shows the difference in the predicted size compared to the original prediction for each sized particulate (5-25 μm). Even for 100 occlusions, the neural network output remains within ~ 0.2-microns of the original prediction. The network only produces an error of greater than 1 micron when 1000 pixels were occluded. Interestingly, the larger the size of the particulate from which the light is scattered, the less the effect of the occluded pixels. This can perhaps be understood by looking at the scattering patterns and the maximum neural network activation of the final ReLu layer (normalised to 1) from the highest error size (8 μm) and the lowest error size (25 μm), see figure 4(b). For 8 μm, the maximum activation (red region), which shows what region is strongly triggering the neural network to categorise a certain number, is much smaller than that for the 25 μm scattering pattern. This indicates that 1000 pixels could have a greater effect on the features that allow for determination of the 8 μm scattering pattern, and thus lead to greater error in determining its size.

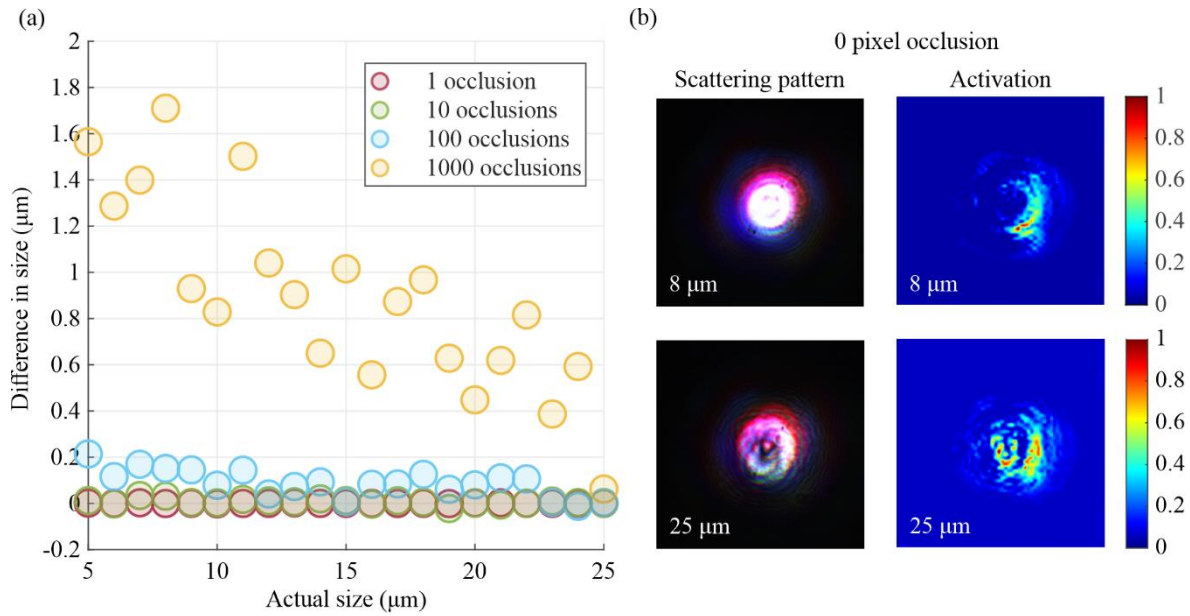

Figure 4. (a) Difference between the prediction of the Y-dimension for particulate scattering patterns without occlusions and scattering patterns containing occlusions, for a total of 1 (red), 10 (green), 100 (blue) and 1000 (yellow) randomly located single occlusion pixels. (b) Scattering patterns and maximum activation of the neural network for 8 µm and 25 µm with zero pixel occlusions.

The results above demonstrate that the network was robust up to the case of 10 single pixel occlusions, and even for 100 single pixel occlusions, the error was ~ 0.2 µm. With the optical arrangement used here, occlusion of a single pixel would require a particle of ~ 15 µm in geometric diameter, however, many particulates, as shown in the SEM images, can be larger. Indeed, smoke and cement dust can be as large as 100s of microns and smaller particulates can also agglomerate in the environment to form larger sensor occlusions. Therefore, if sensing larger particulates (> 15 µm) was desired or agglomeration could not be mitigated inside the detector, understanding how size determination is affected by larger particulates or agglomerates occluding the sensor and whether errors can be mitigated using methods such as data augmentation, would be invaluable.

The size of the occlusions was increased to 5 and 10-pixel diameters on the test images, which were the fed into the neural network. Following this, the training data were augmented to include 1, 5 and 10-pixel diameter occlusions to train a neural network that could perform with greater accuracy on occluded data. Initially, the neural network trained on augmented data was tested on 1-pixel diameter occlusions, in order to determine if the network was suitable for improving on the accuracy of the predictions presented in figure 4. The results are shown in figure 5, and it is evident that the difference between the predicted sizes for particulate scattering patterns without occlusions and with, has decreased across all sizes. For example, the prediction for 5 µm sized particulate has decreased in error from ~ 1.6 µm (figure 4(a)) down to ~ 0.3 µm.

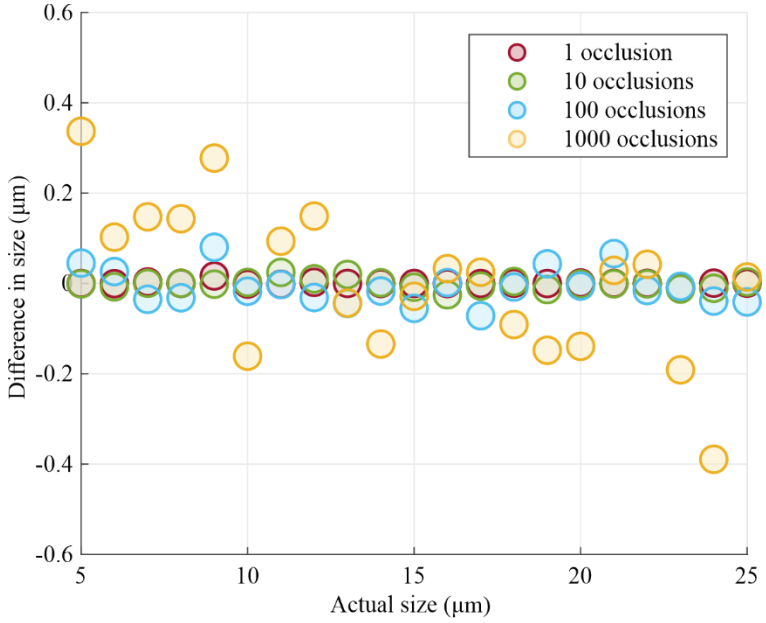

Figure 5. Difference between the prediction, using the neural network trained on augmented data, of the Y-dimension for particulate scattering patterns without occlusions and scattering patterns containing occlusions with single pixel width, for a total of 1 (red), 10 (green), 100 (blue) and 1000 (yellow) randomly located occlusion pixels.

Figure 6(a) shows the absolute difference in the root-mean-square error (RSME) of the prediction of all the test particulates (5-25 µm) for the original neural network and augmented neural network for 1, 5 and 10-pixel diameter occlusion, numbering 1, 10, 100 and 1000, compared with the RMSE of the prediction of all the test scattering patterns with 0 occlusions. It is evident from the plot in figure 6(a) that the network trained with augmented data produced a reduction in the RMSE of the prediction of

size by approximately 0.5 microns for 100 5-pixel diameter occlusions, and at least 1 micron for 10-pixel diameter occlusions totalling 100 and 1000. Figure 6(b) shows examples of scattering patterns occluded with 100 of 1, 5 and 10-pixel diameter circles. For 5-pixel diameter occlusions at numbers totalling 1000 on each pattern, a reduction of an RMSE of over 3 microns was achieved, meaning a potential 62% increase in accuracy could be achieved even when 20% of the sensor's surface is occluded.

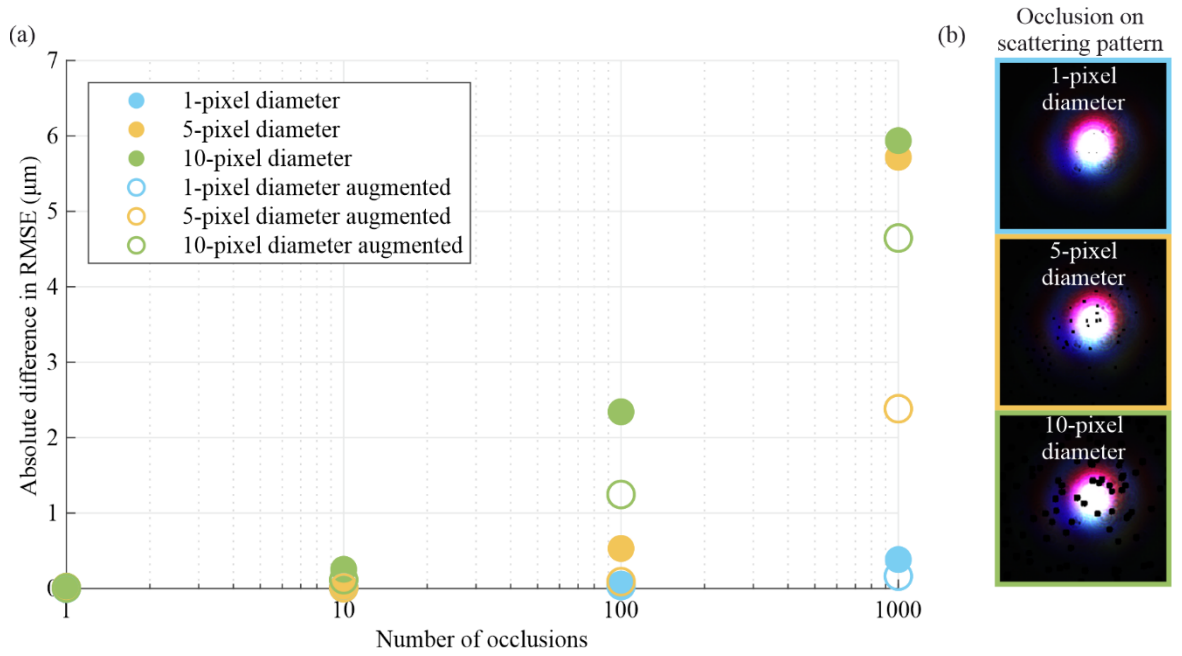

Figure 6. (a) Absolute difference between the RMSE of all the tested scattering patterns (5-25 µm) predictions by neural network trained on original unedited data (solid circles) and a neural network trained on augmented data (hollow circles), compared with the RMSE predictions without occlusions. (b) Scattering patterns showing example occluded scattering patterns covered randomly in 100 zero intensity pixel circle occlusions of 1, 5 and 10-pixel diameters.

The results demonstrate that such a detector could remain accurate for longer via the use of adequate training of a neural network, if the number of particulates deposited and occluding the sensor's surface increases over time. For example, it was found in one study that dry deposition of an upper limit of 64 µm diameter particulates (~ 5-pixels diameter occlusions in our work) onto measurement surfaces in the northern Lake Tahoe Basin (USA) could occur at a flux of ~ 10^6 m^{-2} day^{-1} [47], which in this work, corresponds to ~ 10 $sensor^{-1}$ day^{-1}, where our sensor area is approximately 4 mm^2. Thus, as an example, if such deposition conditions occurred on the detector's sensor surface described in this work, and if a detector's acceptable error from the original Y-dimension was < 3 µm, for 5-pixel diameter occlusions in figure 6(a), this would mean that only 100 occluding particulates could be deposited before the error increases beyond the acceptable limit, which corresponds to ~ 30 days. However, if the training data were augmented, then the detector could last for an extra 900 occluding particulates (~ 240 days), without the need for cleaning the detector's sensor surface.

To further improve the accuracy of the neural network, training with augmented data that included occlusions with non-spherical shapes could further aid in increasing the precision of a detector, since, as shown in the SEM images of urban particulates, such objects can be non-symmetric in shape. In addition, incorporating information about the 3-dimensional structure of the particulates into neural network training could perhaps aid in accuracy of a single dimension, owing to the extra information provided, while using a smaller laser beam could lead to less of the unscattered beam saturating pixels on the camera sensor and allow light scattered from smaller particulates to be adequately recorded.

Conclusion

In conclusion, we have shown that neural networks can determine the size of urban particulate matter from their scattering pattern obtained when the particulates were illuminated using 3 different laser beams. In addition, we have shown a neural network's robustness to single pixel occlusion of the sensor and have shown that when the regions of the scattering pattern become significantly occluded, either via numbers or sizes, augmenting the training data to include such occlusions can improve the accuracy of a neural network in determining the size of particulates. In practice, another algorithm (e.g. summing up all zero intensity pixels) could be used to determine the amount of contamination on the sensor to inform the user to clean such a sensor, in order to aid in avoiding incorrect predictions of particulate matter sizes. The results here demonstrate a key technique for enabling deep learning to be applied to sensing technologies for deployment in the real-world environment.

Acknowledgements

BM was supported by an EPSRC Early Career Fellowship (EP/N03368X/1) and EPSRC grant (EP/T026197/1). ML was supported by a BBSRC Future Leader Fellowship (BB/P011365/1) and a Senior Research Fellowship from the National Institute for Health Research Southampton Biomedical Research Centre. Data supporting this manuscript is available at https://doi.org/10.5258/SOTON/D1668.

ORCID iDs

James A Grant-Jacob https://orcid.org/0000-0002-4270-4247

Matthew Praeger https://orcid.org/0000-0002-5814-6155

Matthew Loxham https://orcid.org/0000-0001-6459-538X

Robert W Eason https://orcid.org/0000-0001-9704-2204

Ben Mills https://orcid.org/0000-0002-1784-1012

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

References

[1] Holgate S T 2017 'Every breath we take: the lifelong impact of air pollution'--a call for action *Clin. Med. (Northfield. Il).* **17** 8–12

[2] Peters R, Ee N, Peters J, Booth A, Mudway I and Anstey K J 2019 Air pollution and dementia: a systematic review *J. Alzheimer's Dis.* **70** S145--S163

[3] Osborne N J, Alcock I, Wheeler B W, Hajat S, Sarran C, Clewlow Y, McInnes R N, Hemming D, White M, Vardoulakis S and Fleming L E 2017 Pollen exposure and hospitalization due to asthma exacerbations: daily time series in a European city *Int. J. Biometeorol.* **61** 1837–48

[4] D'Amato G, Baena-Cagnani C E, Cecchi L, Annesi-Maesano I, Nunes C, Ansotegui I, D'Amato M, Liccardi G, Sofia M and Canonica W G 2013 Climate change, air pollution and extreme events leading to increasing prevalence of allergic respiratory diseases *Multidiscip. Respir. Med.* **8** 12

[5] Sheffield P, Roy A, Wong K and Trasande L 2011 Fine particulate matter pollution linked to respiratory illness in infants and increased hospital costs. *Health Aff.* **30** 871–8

[6] Seaton A, Godden D, MacNee W and Donaldson K 1995 Particulate air pollution and acute health effects *Lancet* **345** 176–8

[7] Europe W R O for 2006 *Air Quality Guidelines Global Update 2005: Particulate Matter, ozone, nitrogen dioxide and sulfur dioxide*

[8] WHO 2016 Global Urban Ambient Air Pollution Database https://www.who.int/airpollution/data/cities-2016/en/ (Accessed 02/03/2021)

[9] Wei Y, Wang Y, Di Q, Choirat C, Wang Y, Koutrakis P, Zanobetti A, Dominici F and Schwartz J D 2019 Short term exposure to fine particulate matter and hospital admission risks and costs in the Medicare population: time stratified, case crossover study *BMJ* **367**

[10] Kelly F J and Fussell J C 2012 Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter *Atmos. Environ.* **60** 504–26

[11] Vieira L de J, Soares T L, Rossi M L, Alves A A C, Santos F de A R dos and Souza F V D 2012 Viability, production and morphology of pollen grains for different species in the genus Manihot (Euphorbiaceae) *Acta Bot. Brasilica* **26** 350–6

[12] McInnes R N, Hemming D, Burgess P, Lyndsay D, Osborne N J, Skjøth C A, Thomas S and Vardoulakis S 2017 Mapping allergenic pollen vegetation in UK to study environmental exposure and human health *Sci. Total Environ.* **599**–**600** 483–99

[13] Gilmour P S, Brown D M, Lindsay T G, Beswick P H, MacNee W and Donaldson K 1996 Adverse health effects of PM10 particles: involvement of iron in generation of hydroxyl radical. *Occup. Environ. Med.* **53** 817–22

[14] Pope C A and Dockery D W 1992 Acute Health Effects of PM10 Pollution on Symptomatic and Asymptomatic Children *Am. Rev. Respir. Dis.* **145** 1123–8

[15] Doiron D, de Hoogh K, Probst-Hensch N, Fortier I, Cai Y, De Matteis S and Hansell A L 2019 Air pollution, lung function and COPD: results from the population-based UK Biobank study *Eur. Respir. J.* **54** 1802140

[16] Liu S, Zhou Y, Liu S, Chen X, Zou W, Zhao D, Li X, Pu J, Huang L, Chen J, Li B, Liu S and Ran P 2017 Association between exposure to ambient particulate matter and chronic obstructive pulmonary disease: results from a cross-sectional study in China *Thorax* **72** 788 LP – 795

[17] Guarnieri M and Balmes J R 2014 Outdoor air pollution and asthma *Lancet* **383** 1581–92

[18] Raaschou-Nielsen O, Andersen Z J, Beelen R, Samoli E, Stafoggia M, Weinmayr G, Hoffmann B, Fischer P, Nieuwenhuijsen M J, Brunekreef B, Xun W W, Katsouyanni K, Dimakopoulou K, Sommar J, Forsberg B, Modig L, Oudin A, Oftedal B, Schwarze P E, Nafstad P, De Faire U, Pedersen N L, Östenson C-G, Fratiglioni L, Penell J, Korek M, Pershagen G, Eriksen K T, Sørensen M, Tjønneland A, Ellermann T, Eeftens M, Peeters P H, Meliefste K, Wang M, Bueno-de-Mesquita B, Key T J, de Hoogh K, Concin H, Nagel G, Vilier A, Grioni S, Krogh V, Tsai M-Y, Ricceri F, Sacerdote C, Galassi C, Migliore E, Ranzi A, Cesaroni G, Badaloni C, Forastiere F, Tamayo I, Amiano P, Dorronsoro M, Trichopoulou A, Bamia C, Vineis P and Hoek G 2013 Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE) *Lancet Oncol.* **14** 813–22

[19] Du Y, Xu X, Chu M, Guo Y and Wang J 2016 Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence *J. Thorac. Dis.* **8** E8–19

[20] Maher B A, Ahmed I A M, Karloukovski V, MacLaren D A, Foulds P G, Allsop D, Mann D M A, Torres-Jardón R and Calderon-Garciduenas L 2016 Magnetite pollution nanoparticles in the human brain *Proc. Natl. Acad. Sci.* **113** 10797 LP – 10801

[21] Calderón-Garcidueñas L, González-Maciel A, Mukherjee P S, Reynoso-Robles R, Pérez-Guillé B, Gayosso-Chávez C, Torres-Jardón R, Cross J V, Ahmed I A M, Karloukovski V V

and Maher B A 2019 Combustion- and friction-derived magnetic air pollution nanoparticles in human hearts *Environ. Res.* **176** 108567

[22] Bové H, Bongaerts E, Slenders E, Bijnens E M, Saenen N D, Gyselaers W, Van Eyken P, Plusquin M, Roeffaers M B J, Ameloot M and Nawrot T S 2019 Ambient black carbon particles reach the fetal side of human placenta *Nat. Commun.* **10** 3866

[23] Loxham M, Davies D E and Holgate S T 2019 The health effects of fine particulate air pollution. *BMJ* **367** l6609

[24] Zhang R, Liu C, Zhou G, Sun J, Liu N, Hsu P-C, Wang H, Qiu Y, Zhao J, Wu T, Zhao W and Cui Y 2018 Morphology and property investigation of primary particulate matter particles from different sources *Nano Res.* **11** 3182–92

[25] Agrawal Y C, Whitmire A, Mikkelsen O A and Pottsmith H C 2008 Light scattering by random shaped particles and consequences on measuring suspended sediments by laser diffraction *J. Geophys. Res. Ocean.* **113**

[26] Gibbs R J 1978 Light scattering from particles of different shapes *J. Geophys. Res. Ocean.* **83** 501–2

[27] Bohren C F and Huffman D R 2008 *Absorption and scattering of light by small particles* (John Wiley & Sons)

[28] Mills B, Chau C F, Rogers E T F, Grant-Jacob J, Stebbings S L, Praeger M, de Paula A M, Froud C A, Chapman R T, Butcher T J, Baumberg J J, Brocklesby W S and Frey J G 2008 Direct measurement of the complex refractive index in the extreme ultraviolet spectral region using diffraction from a nanosphere array *Appl. Phys. Lett.* **93** 231103

[29] Mishchenko M I, Hovenier J W and Travis L D 2000 Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications *Meas. Sci. Technol.* **11** 1827

[30] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A 2015 Going deeper with convolutions *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 1–9

[31] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436

[32] Serre T, Wolf L, Bileschi S, Riesenhuber M and Poggio T 2007 Robust object recognition with cortex-like mechanisms *IEEE Trans. Pattern Anal. Mach. Intell.* 411–26

[33] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in neural information processing systems* pp 1097–105

[34] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition *arXiv Prepr. arXiv1409.1556*

[35] Mollahosseini A, Chan D and Mahoor M H 2016 Going deeper in facial expression recognition using deep neural networks *2016 IEEE winter conference on applications of computer vision (WACV)* pp 1–10

[36] Ranjan R, Sankaranarayanan S, Bansal A, Bodla N, Chen J-C, Patel V M, Castillo C D and Chellappa R 2018 Deep learning for understanding faces: Machines may be just as good, or better, than humans *IEEE Signal Process. Mag.* **35** 66–83

[37] Schofield D, Nagrani A, Zisserman A, Hayashi M, Matsuzawa T, Biro D and Carvalho S 2019 Chimpanzee face recognition from videos in the wild using deep learning *Sci. Adv.* **5** eaaw0736

[38] Grant-Jacob J A, Mackay B S, Baker J A G, Heath D J, Xie Y, Loxham M, Eason R W and

Mills B 2018 Real-time particle pollution sensing using machine learning *Opt. Express* **26** 27237–46

[39] Grant-Jacob J A, Jain S, Xie Y, Mackay B S, McDonnell M D T, Praeger M, Loxham M, Richardson D J, Eason R W and Mills B 2019 Fibre-optic based particle sensing via deep learning *J. Phys. Photonics* **1** 44004

[40] Grant-Jacob J A, Xie Y, Mackay B S, Praeger M, McDonnell M D T, Heath D J, Loxham M, Eason R W and Mills B 2019 Particle and salinity sensing for the marine environment via deep learning using a Raspberry Pi *Environ. Res. Commun.* **1** 035001

[41] Mikołajczyk A and Grochowski M 2018 Data augmentation for improving deep learning in image classification problem *2018 international interdisciplinary PhD workshop (IIPhDW)* pp 117–22

[42] Salamon J and Bello J P 2017 Deep convolutional neural networks and data augmentation for environmental sound classification *IEEE Signal Process. Lett.* **24** 279–83

[43] Ding J, Li X and Gudivada V N 2017 Augmentation and evaluation of training data for deep learning *2017 IEEE International Conference on Big Data (Big Data)* pp 2603–11

[44] Wu T, Tong L and Vorobeychik Y 2019 Defending against physically realizable attacks on image classification *arXiv Prepr. arXiv1909.09552*

[45] Theagarajan R, Chen M, Bhanu B and Zhang J 2019 Shieldnets: Defending against adversarial attacks using probabilistic adversarial robustness *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp 6988–96

[46] Grant-Jacob J A, Praeger M, Loxham M, Eason R and Mills B 2020 Lensless imaging of pollen grains at three-wavelengths using deep learning *Environ. Res. Commun.*

[47] Tai A Y-C, Chen L-W A, Wang X, Chow J C and Watson J G 2017 Atmospheric deposition of particles at a sensitive alpine lake: Size-segregated daily and annual fluxes from passive sampling techniques *Sci. Total Environ.* **579** 1736–44