

TOPICAL REVIEW • OPEN ACCESS

## Deep learning in airborne particulate matter sensing: a review

To cite this article: James A Grant-Jacob and Ben Mills 2022 *J. Phys. Commun.* **6** 122001

View the [article online](#) for updates and enhancements.

### You may also like

- [Analysis and Prevention of Urban River Pollution](#)  
Yibo Wang, Yingxin Liu, Jinming Huang et al.
- [The estimation of the load of non-point source nitrogen and phosphorus based on observation experiments and export coefficient method in Three Gorges Reservoir Area](#)  
X X Tong, B Hu, W S Xu et al.
- [Application TAPM-AERMOD system model to study impacts of thermal power plants in SouthEast and SouthWest areas to the air quality of HCMC: current status and according to Vietnam power planning VII toward 2030](#)  
Khue Hoang Ngoc Vu, Hang Thi Thuy Nguyen, Tam Thoai Nguyen et al.



## TOPICAL REVIEW

## Deep learning in airborne particulate matter sensing: a review

## OPEN ACCESS

RECEIVED  
28 August 2022

REVISED  
22 October 2022

ACCEPTED FOR PUBLICATION  
18 November 2022

PUBLISHED  
5 December 2022

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



James A Grant-Jacob\* and Ben Mills

Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, United Kingdom

\* Author to whom any correspondence should be addressed.

E-mail: [J.A.Grant-Jacob@soton.ac.uk](mailto:J.A.Grant-Jacob@soton.ac.uk)

**Keywords:** deep learning, particulate matter, air pollution, sensing

### Abstract

Airborne particulate matter pollution is a global health problem that affects people from all demographics. To reduce the impact of such pollution and enable mitigation and policy planning, quantifying individuals' exposure to pollution is necessary. To achieve this, effective monitoring of airborne particulates is required, through monitoring of pollution hotspots and sources. Furthermore, since pollution is a global problem, which varies from urban areas to city centres, industrial facilities to inside homes, a variety of sensors might be needed. Current sensing techniques either lack species resolution on a world scale, lack real-time capabilities, or are too expensive or too large for mass deployment. However, recent work using deep learning techniques has expanded the capability of current sensors and allowed the development of new techniques that have the potential for worldwide, species specific, real-time monitoring. Here, it is proposed how deep learning can enable sensor design for the development of small, low-cost sensors for real-time monitoring of particulate matter pollution, whilst unlocking the capability for predicting future particulate events and health inference from particulates, for both individuals and the environment in general.

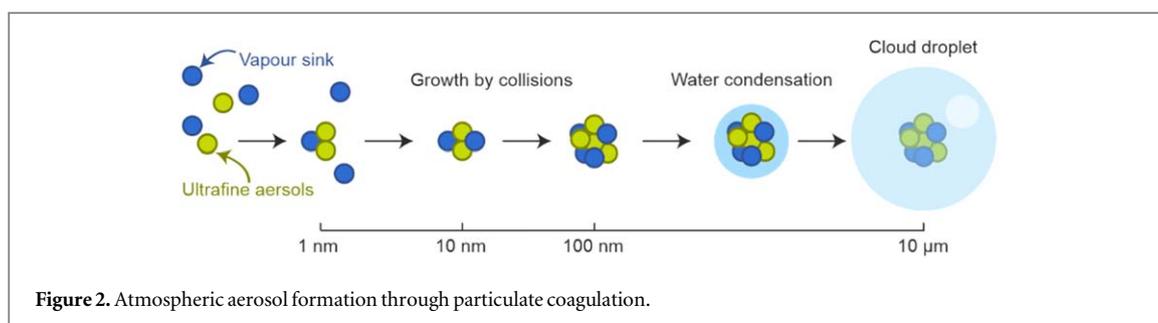
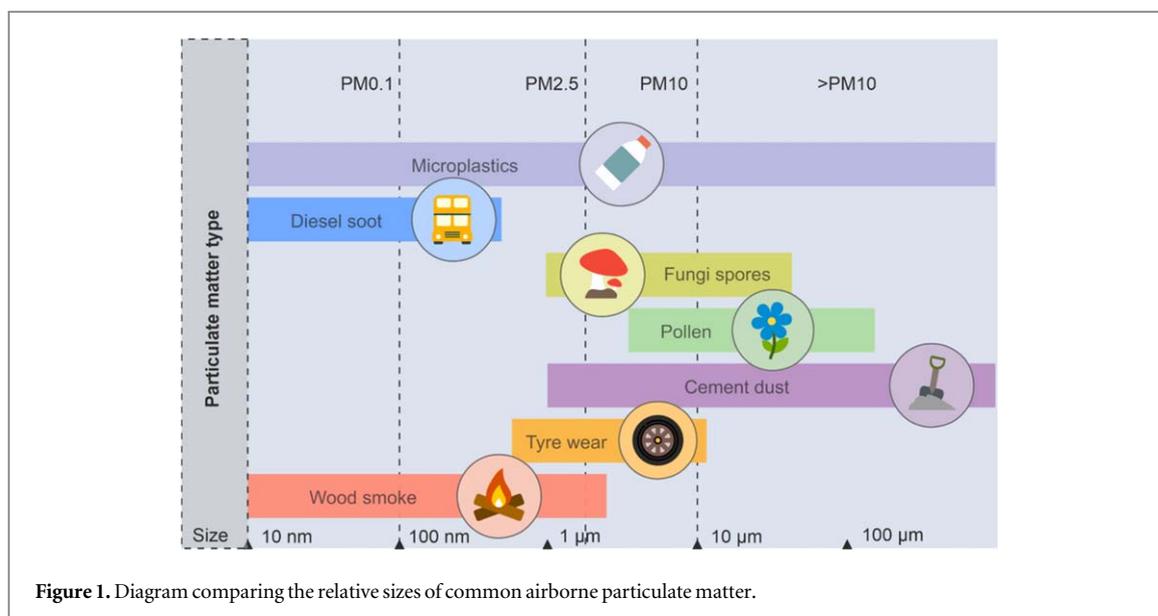
## 1. Introduction

Airborne particulate matter pollution is a worldwide health and environmental problem. To determine sources to reduce their impact, mitigate toxic levels, and predict dangerous episodes (just as current meteorological centres do regarding heatwaves, hurricanes and flooding), precise sensing capability is necessary on an individual, local and global scale. Owing to recent significant advancements in the field of deep learning, this review aims to discuss how such a field has been applied to the domain of particulate matter detection. The review begins by describing the various sources of particulate matter pollution, along with the dynamics of the particulates and the associated health and environmental implications. Following this, in section 3, we discuss the current common methods used for the detecting and analysis of particulate matter present in the atmosphere. Deep learning is then introduced in section 4, where we discuss its history, followed by its application to the field of airborne particulate matter pollution. Finally, in section 5, we discuss the future perspectives of deep learning for particulate matter sensing and propose future ideas that could enable the development of more accurate sensors and stronger synergy between detection, forecasting and inference.

## 2. Airborne particulate matter

### 2.1. Sources

Air pollution can consist of particulate matter in the form of particles that can range from over  $10\ \mu\text{m}$  ( $>\text{PM}_{10}$ ), to below  $10\ \mu\text{m}$  ( $\text{PM}_{10}$ ), below  $2.5\ \mu\text{m}$  ( $\text{PM}_{2.5}$ ), and below  $0.1\ \mu\text{m}$  ( $\text{PM}_{0.1}$ ), where the particles themselves can be formed from a range of chemicals [1] that are present worldwide [2]. Since particulates of all sizes exist in a variety of shapes (symmetrical, asymmetrical, and irregular), the size of a particulate is generally defined by its aerodynamic diameter, which is a diameter equal to a spherical particulate with a density of  $1\ \text{g cm}^{-3}$ , whose aerodynamic behaviour (settling velocity) is equivalent to the particulate in question [3].



Particulate matter emanates from a variety of sources. A significant source of airborne particulate matter is road transport, particularly from the combustion of diesel from public and private vehicles [4]. These particulates are generally carbonaceous materials, including black carbon (carbon-based micron and nano-sized spheres formed from the incomplete combustion of organic matter), elemental carbon, and a range of polycyclic aromatic hydrocarbons (PAHs) [5]. Brake and tyre wear can also produce rubber and metallic particulates [6]. The burning of fossil fuels [7], biofuels [8] and wood [9], also contribute to black carbon particulate matter. Other types of particulate matter include iron and sulphur from steel works [10], metals from mining [11] and cement dust from factories [12]. Particulate matter can also originate from natural sources such as desert dust, whilst sea salt can also contribute to overall atmospheric particulate levels [13]. Outdoor natural biological contributions include pollen from trees, grass and weeds, and spores from fungi [14].

Airborne particulate matter pollution can also occur indoors from a variety of sources such as cooking [15], as well as spray products (e.g., surface cleaners, deodorants and air fresheners [16]), wood burning stoves [17], cigarette smoking and incense [18], fungi [19], as well as general sweeping and cleaning, which can project particulate matter into the air [20]. Additional sources include microplastic fibres from the wear and tear of textiles [21] and asbestos fibres from walls and roofing [22]. Figure 1 shows a diagram summarising common particulates and their size range.

## 2.2. Dynamics

The size and composition of particulate matter, as well as metrological factors, can affect their number density and toxicity [23], and therefore sensing devices should also include the capability for identification of these factors. For example, particulates exist in a variety of sizes, from single spheres and cuboids to shards and agglomerates [24]. Ultrafine particulate matter can evolve and become larger by coagulation via collisions with other particles (see figure 2), by chemical reaction, and by activation in the presence of water vapour in the troposphere [25]. In addition, pollen grains, which usually have a size of 10–100  $\mu\text{m}$ , have been found to have urban particulate matter attached to them [26].

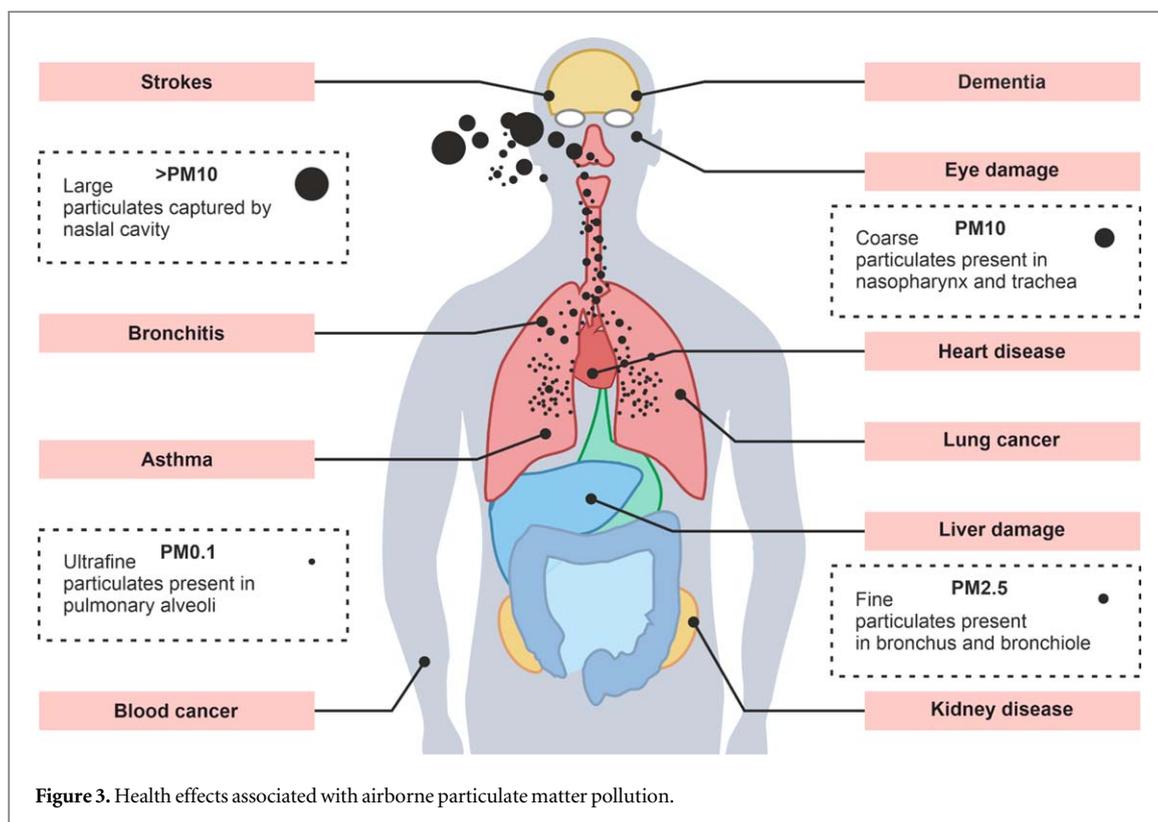


Figure 3. Health effects associated with airborne particulate matter pollution.

Particulate matter can be present at different heights (based on the initial altitude of the emission, air flow, and weather conditions), with concentrations varying depending on meteorological factors and population density [27]. Due to the size and weight of airborne particulates, such pollution can travel in the atmosphere from one area to another [28]. Not only do particulate pollution levels vary from region to region [29], but particulate matter levels can also even vary over hours [30], months [31], and on a yearly basis [32]. Therefore, particulate sensors would need to be able to detect over both short and long time periods, at different places and different altitudes.

### 2.3. Health

Owing to the small size, shape and variety of chemical compositions, airborne particulate matter directly results in a range of health conditions [33, 34] and is estimated to contribute to around nine million deaths globally per year [35, 36]. More specifically, fine particulate matter can have a detrimental effect on life expectancy, as documented in a study in the United States of America [37].

Figure 3 presents a diagram showing the pathway of different sized particulates into the human body and highlights associated health impacts of such particulates. The eyes are particularly susceptible to dust and other particulates, and PM2.5 has been found to trigger ocular hypertension [38]. Particulates on the size scale of 10–100  $\mu\text{m}$ , such as pollen grains, can get trapped in the nasal cavity and can lead to allergic rhinitis, with approximately 10% to 30% of adults around the world suffering from the condition [39].

In general, smaller particulates are more likely to travel further along the respiratory tract and into the lungs, and larger particulate (PM10 and >PM10) are instead generally found more frequently in the upper airways. Indeed, cement dust exposure can enter the respiratory tract and has been linked to acute respiratory illnesses such as sneezing, coughing and shortness of breath [40]. Microplastics (PM10), such as polyethylene and polyethylene terephthalate fibres, have been found in tissue from the upper, middle, and lower part of the human lungs [41].

Cigarette smoke can contain considerable amount of PM2.5, and high levels have been associated with lung cancer mortality [42], and PM2.5 urban particulate matter has been linked to asthma exacerbation [43]. Since PM2.5 includes particles below 100 nm, such particulates can enter alveolar, allowing them to enter the blood circulation system and in turn travel to various organs of the human body. There is also evidence that particulates could lead to liver disease [44], and long-term exposure has demonstrated an increased risk of obtaining non-Hodgkin lymphoma [45]. Further to this, there is indication that particulates accumulating in the brain [46] are associated with Alzheimer's disease [47], dementia [48], strokes [49] and perhaps autism [50].

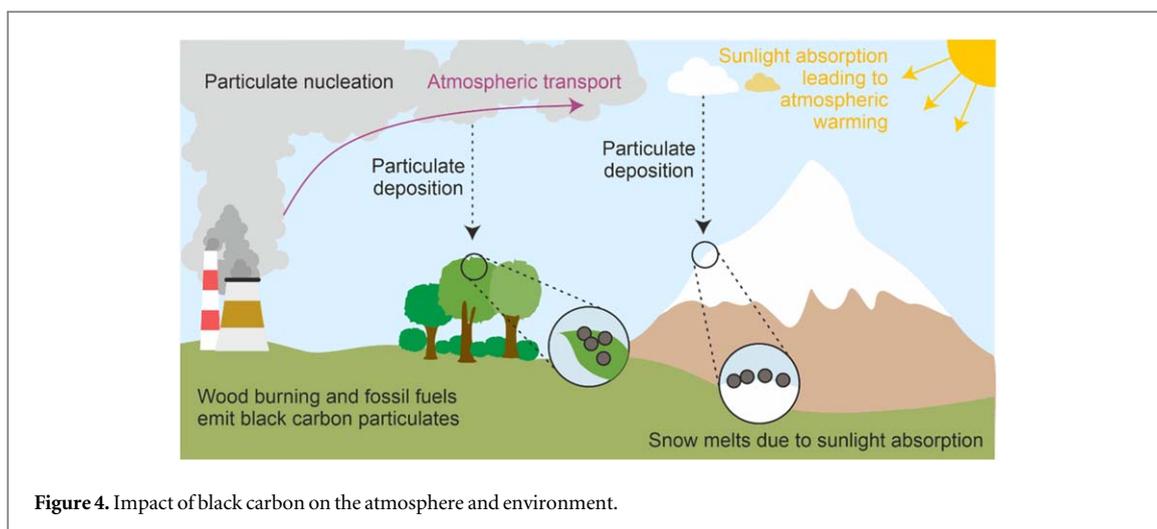


Figure 4. Impact of black carbon on the atmosphere and environment.

## 2.4. Environmental effects

Airborne particulates have adverse effects on the environment and contribute to climate change [51]. As shown in the diagram of figure 4, particulate matter such as black carbon from wood and fossil fuel burning can be emitted into the atmosphere and transported via air currents to different areas. Black carbon present in the atmosphere can absorb infrared radiation, and so can lead to heating of the atmosphere [52]. In addition, atmospheric black carbon can deposit on snow, leading to increased melting due to the lowering of snow grain albedo [53].

Particulate matter can deposit on the leaves of vegetation and in the surrounding soil [54]. Metals can be absorbed by the roots [55] and be present on plants, and in turn affect the biochemical processes [56]. The direct deposition of particulate matter onto the surfaces of leaves can cause damage due to pH of the particulates [57], and potentially inhibit photosynthesis and respiration leading to the decaying of leaf matter, with the associated negative effects on agriculture [58].

## 3. Current sensing methods and monitoring

The World Health Organisation (WHO) has set limits on the amount of particulate matter that is safe, with values of fine particulate matter (PM<sub>2.5</sub>) of  $5 \mu\text{g m}^{-3}$  annual mean with  $15 \mu\text{g m}^{-3}$  24 h mean, and coarse particulate matter (PM<sub>10</sub>) of  $15 \mu\text{g m}^{-3}$  annual mean with  $45 \mu\text{g m}^{-3}$  24 h mean [59]. Since particulate levels often exceed such values throughout the world [60], being able to accurately measure levels is vital for human health and understanding the sources of particulate matter. As discussed in this section, a variety of sensing methods exist, each with their own advantages and disadvantages.

### 3.1. Gravimetric

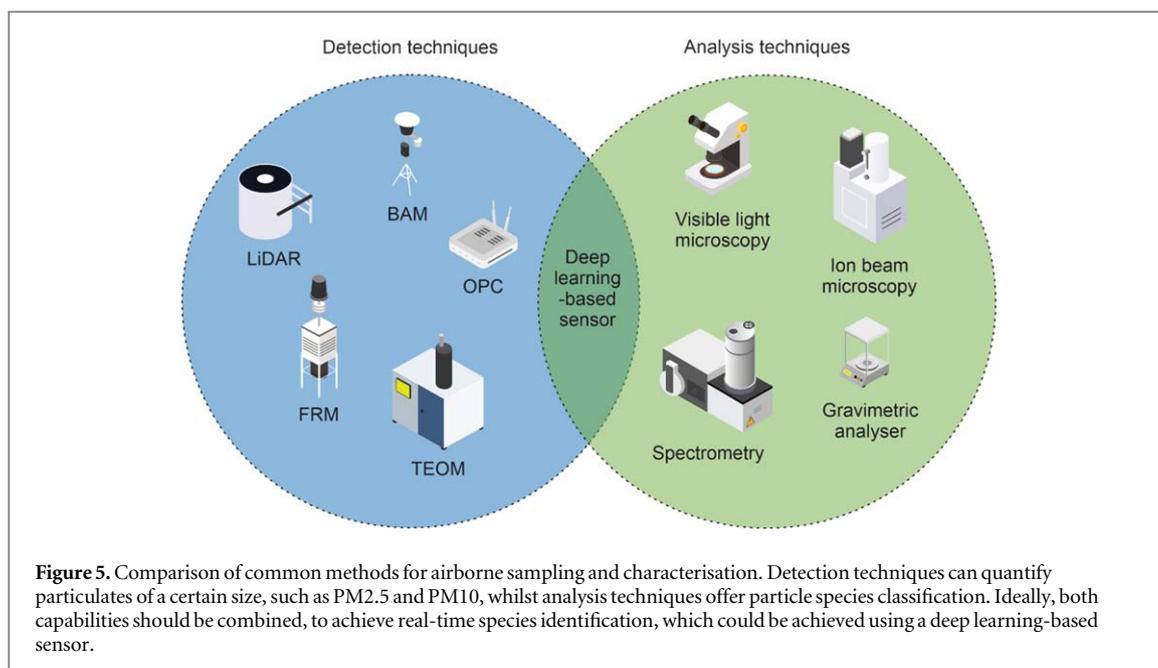
A critical capability of a particulate sensor is to be able to determine or infer masses and concentration of particulates. For both the PM<sub>10</sub> and PM<sub>2.5</sub> standards, the Federal Reference Method (FRM), as defined by the U.S. Environmental Protection Agency (EPA), states that the reference technology used for measuring PM<sub>10</sub> and PM<sub>2.5</sub> should be based on gravimetric analysis of particulates collected over a 24 h time period [61]. Whilst measurements can be sorted into PM<sub>10</sub> and PM<sub>2.5</sub>, species specific measurements cannot generally be obtained, and this technique lacks real-time capability.

### 3.2. Oscillation

Tapered element oscillating microbalance (TEOM) sensors consist of a filter on the end of a glass tip whose vibration frequency is measured electronically [62, 63]. As particulate matter is deposited onto the filter, the glass tip oscillation frequency changes. The particulate matter mass can then be directly inferred as the oscillation frequency will decrease as particulate mass is added. Whilst such measurements can be very precise, species identification is not possible.

### 3.3. Radiation

A beta attenuation monitor (BAM) utilises  $\beta$ -ray attenuation to quantify the particulate matter concentration in air [64]. Air is passed through an inlet and particulate matter is impacted onto a glass fibre tape. The tape is situated between a  $\beta$ -radiation source and a scintillator, so that matter collected on the surface of the tape



attenuates the radiation and therefore affects the signal on the scintillator. The amount of attenuation is generally proportional to the amount of particulate matter sampled.

### 3.4. Optical

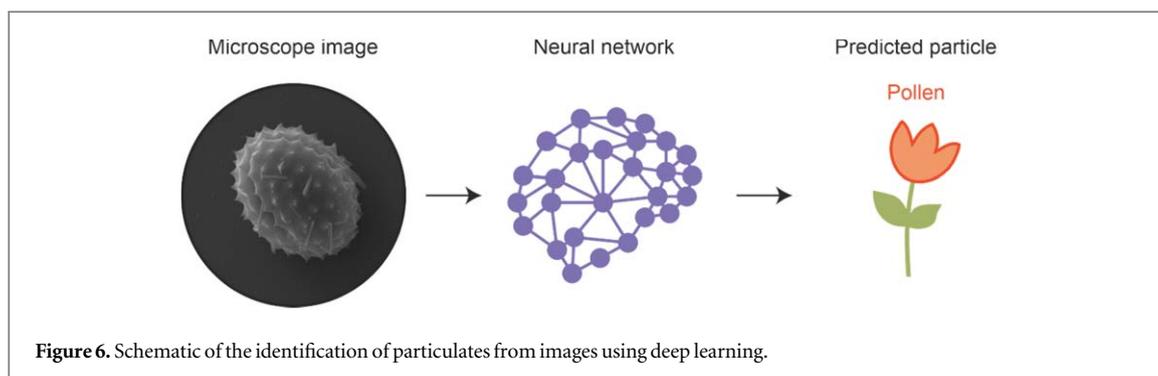
Satellite based sensing provides the advantage of large area and volumetric sensing. Columnar satellite-derived aerosol optical thickness (AOT) values can be related, via linear regression, to surface PM<sub>2.5</sub> mass measurements, based on EPA guidelines [65]. Engel-Cox *et al* outlined the use of Light Detection and Ranging (LiDAR) at 532 nm wavelength in combination with aerosol optical depth (AOD) measurements and ground-based PM<sub>2.5</sub> monitoring data to create 3D mapping of PM<sub>2.5</sub> mass concentration [66].

Since the optical properties of particulates are dependent on their chemical and structural properties, analysis of the light scatter can be used to determine particulate levels, as in the case of optical particle counters (OPCs) [67]. OPCs generally use laser or LED light to illuminate particulates in a specified volume of air and capture the scattered light using a photodiode detector. Whilst these optical scattering sensors are relatively simple, small and cheap, and allow for real-time monitoring by scientists and communities, they currently lack the capability for species specific identification [68].

### 3.5. Spore traps

Since particulate matter exists in a variety of shapes, sizes, and composition, with their associated effects dependent on these parameters, there is a clear importance for identification of each individual particulate. Volumetric air sampling traps, such as the Burkard Spore Trap, can be used to collect particulate matter of a variety of sizes to enable analysis of airborne particulates over a 24 h or 7 day period [69]. These traps are approximately 1 m tall, and air is drawn in through an inlet such that particulates are impacted onto an adhesive surface such as tape. Following a desired period of monitoring, the tape is removed from the trap and inspected in a laboratory via methods such as visible light microscopy [70], electron microscopy [71] and x-ray photoelectron spectroscopy (XPS) [72], to allow for particulate species identification. Since species identification occurs post sample collection in a laboratory, spore traps do not enable real-time measurements.

As such, alternative methods are required for more immediate mitigation. Ideally, a sensor should be able to determine the species of individual particulate matter, as well as their size and shape, in real-time, and with a low cost and small footprint to enable mass deployment. Figure 5 presents a Venn diagram of common existing detection and analysis technique, with a deep learning-based sensor in the overlap region, since, as will be described in the next section, deep learning excels at using large data to improve the capability of current techniques, and to synergise techniques.



## 4. Deep learning

### 4.1. History

Deep learning is a subset of artificial intelligence (AI) that uses neural network algorithms to process information [73]. The information that a neural network receives is passed through multiple layers, where the output of a layer acts as the input to the next. Deep learning can utilise convolutional filters to perform feature extraction, to enable a neural network to understand similarity and trends in data. Deep learning has its origins in 1943 when the first computer model using threshold logic to mimic the thought process was developed by Warren McCulloch and Walter Pitts [74].

Whilst algorithms were being developed and improved, it was the advent of mass market graphics processing units (GPUs) for desktop computers [75] that allowed for parallel computing, hence allowing complex problems to be separated into thousands or millions of separate tasks and processed simultaneously, and in doing so, unlock huge capabilities for deep learning algorithms. By 2012, GPUs had a significant advantage over CPUs for processing images, with  $4 \times$  in GFLOPS (giga floating point operations per second) processing in some instances [76], meaning that deep learning algorithms were routinely performed on computers with dedicated GPUs for training. Indeed, ImageNet's classification with deep convolutional neural networks (CNN) was demonstrated by Krizhevsky *et al* [77], in which they were able to classify 1.2 million images into 1000 categories (leopard, container ship etc.) with a top-5 error rate of 17%. The results were significantly better than previous work and used two NVIDIA 580 3GB GPUs, taking 6 days to accomplish.

In 2014, generative adversarial networks (GANs) were developed [78], which allow for realistic image generation of written digits, and photographs of faces and animals, while a conditional generative adversarial network (cGAN) was implemented by Isola *et al* [79] for image transformation from one domain to another, such as black and white photographs to colour photographs. Neural networks have since been applied to almost all scientific domains, including the natural world, with examples including bird sound identification [80], wild chimpanzee face recognition [81] and orca sound detection [82].

### 4.2. Deep learning applications to airborne pollution sensing

Due to the capability of deep learning for pattern recognition, such a technique has been implemented in the categorisation of images of airborne particulates, as well as their light scattering patterns, holographic images, and fluorescence signature. Work has been shown to enable the identification of particulates beyond existing techniques such as manual identification of pollen in microscope images. In addition, owing to the ability of neural networks in processing multidimensional data, neural networks have also shown promise in the ability to forecast pollution levels based on limited amount of data. As we discuss in the following subsections, deep learning has been implemented successfully in a variety of areas, from sensing to forecasting.

#### 4.2.1. Optoelectronics

Being able to automatically identify particles from their images is important for classification and source apportionment. As shown in the schematic in figure 6, a trained neural network can be used to identify a particulate, in this case a pollen grain, from a microscope image. As discussed in this section, deep learning has been applied extensively to the field of optics.

Daoud *et al* trained a CNN to classify pollen grain species directly from visible light microscope and scanning electron microscope (SEM) images [83]. When testing on 30 different species, the method achieved  $\sim 90\%$  classification rate for visible light microscope images and  $\sim 94\%$  classification rate when testing on SEM images. The CNNs were more accurate at classification than traditional machine learning methods, such as support vector machine (SVM) classifiers [84] trained on features extracted using fractal dimension, grey level co-occurrence matrix, Gabor features, histograms of oriented gradient descriptors, and local binary pattern

histogram, which achieved classification accuracies of  $\sim 72\%$ ,  $\sim 51\%$ ,  $\sim 67\%$ ,  $\sim 62\%$  and  $\sim 77\%$ , respectively when testing on light microscope images. Likewise, when testing on SEM images, these classifiers achieved lower accuracy than the CNN, with values of  $\sim 61\%$ ,  $\sim 48\%$ ,  $\sim 60\%$ ,  $\sim 50\%$  and  $\sim 72\%$ , respectively.

Classifying only visible light microscopy using AlexNET CNN [77] has shown to achieve an accuracy of 98% for 46 different classes of pollen grain [85]. This result is much higher than manual classification processes, such as that used by the medical and pharmaceutical industry, which were reported to be  $\sim 67\%$ , and higher than classification via an SVM of  $\sim 64\%$  used by others to classify even fewer pollen grains [86].

Using data from an automatic trap system (Bio-Aerosol Analyzer BAA500 [87]), which records images of particulates when they have been captured, the Inception v3 CNN [88] was used to enhance the capability of the BAA500 algorithm and accurately classify 31 taxa of pollen [89].

Other optical methods combined with deep learning include holography, in which a mobile microscopic holography device was developed. In this method, particulates were impacted onto a sticky substrate to allow for holographic imaging and then CNNs were used to reconstruct images of pollen (i.e., transform the holographic image into the real image) and classify 5 different species of pollen as well as dust [90]. The device was also tested in the real-world on oak tree pollen. The device was able to process bioaerosols at a throughput of  $13 \text{ l min}^{-1}$ , and was tested on *Aspergillus* spore, *Alternaria* spore, Bermuda grass pollen, oak tree pollen and ragweed pollen, achieving a classification accuracy greater than 94%.

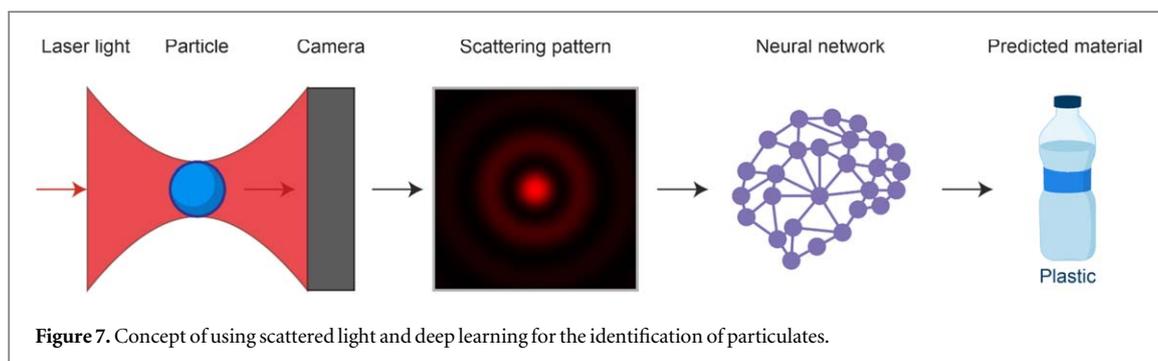
Categorisation of particulate shape is important in understanding their source and potential effects. Yin *et al* used a CNN model with an attention mechanism to identify atmospheric particles of four different structural types, namely flocculent particles, fibrous particles, mineral particles, and spherical particles directly from SEM images [91]. The attention-CNN model was found to have higher classification accuracy compared with CNN and SVM models, since features of the particle can be focused on, instead of non-particle features such as the background.

Using a microfluidics chip to flow pollen grains through a microfluidic device at a rate of 150 grains per second, an AlexNet-based CNN was used to process and classify pollen grains based on their optical images and the output was combined with classified electrical signals in a multimodal approach [92]. The experiment involved classification of eight different classes of pollen, giving an accuracy of 82.8% for the standalone electrical classifier, 84.1% for the standalone optical classifier, and 88.3% for the multimodal approach (processing of combined electrical and optical outputs). It was suggested that the method could also be used to classify other particulates between  $10 \mu\text{m}$  and  $100 \mu\text{m}$ .

A quasi-real-time PM monitoring device was created by integrating smartphone-based digital holographic microscopy via the use of small optics and a laser diode [93]. Holographic speckle images of particulates were obtained and transformed into particulates images via a deep autoencoder (DAE), and regression layers were used to extract important features and predict PM10 and PM2.5 concentrations. Their technique could successfully estimate PM10 and PM2.5 concentrations of dust particulates from the holographic speckle patterns, with high throughput (1.57 s per 100 holograms). The technique was also very accurate, with relative errors of the test datasets for PM10 concentrations being  $11.23\% \pm 9.32\%$ , whilst the relative errors of the test datasets for the PM2.5 concentrations were  $5.81\% \pm 4.46\%$ . Such mobile-optics integration is a potential future direction for particulate sensing, especially for hazardous environments, or for alerting high pollen counts or dangerous toxicity levels.

Whilst current optical particle counters involve using a photodiode detector, using a complementary metal-oxide-semiconductor (CMOS) sensor or charged coupled device (CCD) detector array instead potentially allows more information about the object that scattered the light. As such, by capturing more of the scattered light and thus a 2D spatial scattering pattern, identification of different particulates can be achieved. The concept is presented in figure 7, which shows how laser light is used to illuminate a plastic microparticle, and its scattering pattern is captured by a camera and a neural network identifies the material from the pattern. This 'lensless' sensing method involves capturing the light scattered from the particulate using a camera sensor, rather than collecting the light with an objective lens that would otherwise enable imaging of the object onto a camera sensor. This allows for a smaller, cost-effective sensor design. Using this approach, wood ash and pollen were identified directly from their scattering patterns using a CNN, when the particulates were illuminated via a laser diode [94]. Scattering patterns from each particulate, captured using a camera sensor, were fed into the neural network, taking less than 50 milliseconds to output a predicted category. The neural network was found to correctly identify 43 out of 50 particulates. The results were presented in a confusion matrix, consisting of true positives and true negatives as well as false positives and false negatives. The Matthews correlation coefficient (multiclass classification quality [95]) was calculated to be 0.81 (maximum potential value of 1).

The prediction of pollen grains from scattered light captured via back reflection through an optical fibre has also been demonstrated by the same group [96]. More specifically, pollen grains were illuminated with a single-core fibre and the subsequent scattering patterns were collected via a 30-core fibre, with the scattered light then directed onto a camera sensor for capture. A CNN was employed to categorise 3 different types of pollen grains



using their associated captured scattering patterns. The neural network was able to correctly predict the pollen species with an accuracy of  $\sim 97\%$ . The neural network was also found to be robust to different levels of background light. In addition, the distance between the pollen grains and the end of the fibre was also included as training data for a CNN and when tested, the neural network was able to predict the distance with an associated error of  $\pm 6 \mu\text{m}$ .

To further quantify the particulates that are scattering the light, the scattering signal from particulates in optical scattering can be transformed into an image. Lensless imaging of pollen grains was demonstrated using three lasers of different wavelengths to transform a scattering pattern directly into an image of pollen using a cGAN [97]. Image transformation has also been used to transform images of pollen from one modality (image style) into another, such as from visible light microscope into SEM image, to aid in identification using SEM databases [98]. Further to sensing and imaging, recent work has been carried out using image transformation from a pollen grain in dehydrated state to a pollen grain in a hydrated state, to estimate the original size and shape of the pollen grains prior to dehydration [99]. This could be useful in understanding pollen behaviour and the environment from which it has existed.

#### 4.2.2. Laser-induced spectroscopy

Laser-induced spectroscopy is another technique that has been used in research for particulate matter identification. Since different materials can fluoresce with different wavelengths upon laser illumination, their spectra can be used to identify particulate matter. Using the Rapid-E automatic particle detector to collect scattering and fluorescence data from 29 types of bioaerosols such as pollen and spores, Daunys *et al* [100] found that for clustering unidentified particulates, clustering of fluorescent data played a more important role than clustering scattering data, most likely due to the effect of dehydration on the shape of pollen grains. More specifically, based on fluorescence properties, it was found that *Betulaceae* genus pollen data could be grouped into the same clusters, whilst it was also possible to differentiate *Secale* and *Dactylis* genera from the same family (*Poaceae*) using fluorescent clustering. Further to this work, Tešendić *et al* [101] describe incorporating deep learning into the Rapid-E for real-time monitoring and automatic airborne detection of pollen. They developed a system called 'RealForAll' that retrieves optical data (scattering pattern, fluorescence spectrum and fluorescence lifetime) from Rapid-E devices and classifies the pollen using a CNN. The system provides hourly measurements with a time delay of 15 min (due to batch data processing), via a web-based applications as well as Android and Apple mobile-based applications. Their results showed it was possible to classify 11 pollen types with a Pearson correlation coefficient (a measure of linear association between two sets of data) of  $>0.7$ .

Deep learning has also been used for identifying PM<sub>2.5</sub> source from wood burning, diesel exhaust, and cigarettes using excitation emission matrix (EEM) (excitation wavelength versus emission wavelength versus fluorescence intensity) fluorescence spectroscopy [102]. By collecting material using filters, obtained spectra was used as training data for a CNN to predict the types with an overall accuracy of 89%. To scale to other sources and real-world samples, more data was suggested to be needed.

Relevant to bioaerosol pollution, Hasti *et al* [103] developed a deep learning-based denoiser and spray droplet location predictor to enable accurate estimation of the location of gas turbine combustion spray droplets in light scattered images. A modified CNN (no downsampling to maintain image size), modified ResNet [104] (no downsampling), and modified U-Net were used to denoising the images, with the results showing that the U-Net architecture gave the lowest mean-square error (MSE), with a value of 0.0053 on the validation dataset. Their trained denoising algorithm took  $\sim 2.13$  s per image using a single NVIDIA T4 GPU. Droplet centre prediction was carried out using a CNN-based architecture, giving an average error of 1.35 pixels for the x-pixel location and 0.927 pixels for the y-pixel location of the droplets.

#### 4.2.3. Remote sensing

Satellite-based sensing offers large area monitoring of particulates, and much work has been done in improving the capability using deep learning. Whilst most studies using satellite-derived aerosol optical depth (AOD) PM measurements involve low spatial resolution, a recent study by Imani used deep learning to determine PM<sub>2.5</sub> and PM<sub>10</sub> concentrations using Moderate Resolution Imaging Spectroradiometer (MODIS) satellite images [105]. The work discusses using an LSTM-based neural network that was trained on ground scenes imaged by MODIS, to relate the intensity values of the satellite image bands and particulate matter measurements at different spatial locations. By using publicly available satellite images, the resultant generated PM<sub>2.5</sub> map and PM<sub>10</sub> map of Tehran city showed that the proposed technique is a potentially suitable for simple low-cost PM mapping. The deep learning results are compared with a random forest model employed on the same data, and it was found that in the majority of cases, the deep learning model achieved a lower root-mean-square error (RMSE) and mean absolute percentage error (MAPE), and a higher index of agreement (IA) (difference between observed and estimated values). The model was also compared with state-of-the-art air pollution studies that employ AOD products for particulate matter concentrations estimations. Again, in most cases the results showed lower RMSE. However, it was noted that because the satellite data is affected by cloud levels, the model may not be as suitable when the sky is cloudy.

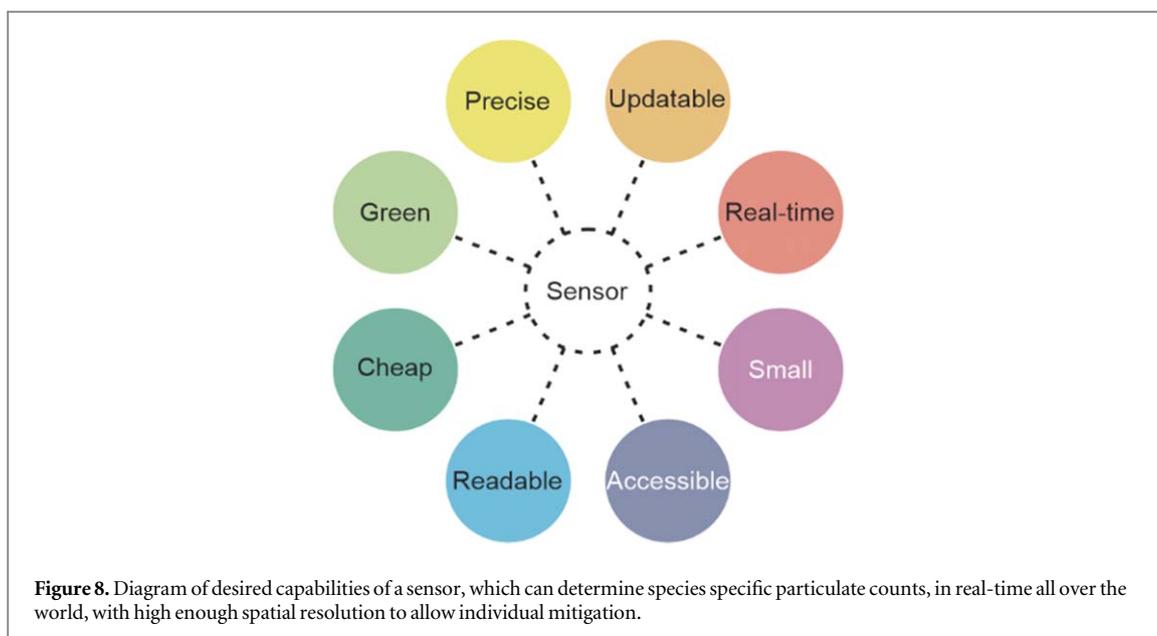
Li *et al* also used deep learning to fuse satellite data from MODIS and meteorological data to estimate ground-level PM<sub>2.5</sub> over large geographical regions in China [106]. They implemented a layer-by-layer pretraining technique along with geographical correlation to improve the accuracy of the PM<sub>2.5</sub> estimation compared with traditional artificial neural network approaches, achieving a cross-validation coefficient of determination (R<sup>2</sup>) value being approximately twice as high at 0.88, and a RMSE value being approximately half as much at 13.03  $\mu\text{g m}^{-3}$  in some cases. From the results, they estimated that more than 80% of the Chinese population live in regions that have an annual mean PM<sub>2.5</sub> greater than the 2015 WHO recommended safe levels.

Sun *et al* [107] also used large deep neural networks to train on large-scale satellite data and meteorological data to identify their spatiotemporal relationships with PM<sub>2.5</sub> particulates to enable ground-based estimations of hourly PM<sub>2.5</sub> levels with 1 km spatial resolution in China. Compared with other state-of-the-art methods, their model achieved better performance in terms of testing in the cold and warm seasons, giving an R<sup>2</sup> of 0.84, RMSE of 19.9  $\mu\text{g m}^{-3}$  and MPE of 11.9  $\mu\text{g m}^{-3}$ , whereas a linear mixed effective (LME) model gave an R<sup>2</sup> of 0.63, RMSE of 29.0  $\mu\text{g m}^{-3}$  and MPE of 18.1  $\mu\text{g m}^{-3}$ , and a support vector regression (SVR) model gave an R<sup>2</sup> of 0.77, RMSE of 21.5  $\mu\text{g m}^{-3}$  and MPE of 12.3  $\mu\text{g m}^{-3}$ .

Yan *et al* have used deep learning to obtain large-scale PM<sub>2.5</sub> concentrations from Himawari-8, a geostationary satellite, to provide high temporal resolution data (hourly intervals) across mainland China [108]. They developed a neural network that they call 'EntityDenseNet', which in contrast to more classical machine learning methods (e.g., backpropagation neural network, extreme gradient boosting, light gradient boosting machine, and random forest), was able to extract spatial and temporal characteristics of PM<sub>2.5</sub> concentrations at a higher accuracy. This work concluded that the coastal city of Tianjin was influenced by air pollution from Hebei, and that PM<sub>2.5</sub> concentrations were closely related in groups of 3 months. Lee *et al* proposed using deep learning for monitoring the hourly spatial concentrations of PM<sub>2.5</sub> over the Korean Peninsula via using geostationary ocean colour imager (GOCI) satellite data, which consisted of multispectral images of top of atmosphere (TOA) reflectance [109]. When comparing the predicted PM<sub>2.5</sub> concentrations with the ground measurements of PM<sub>2.5</sub>, their model was more accurate than the traditional machine learning methods such as random forest, giving an RMSE of 7.042  $\mu\text{g m}^{-3}$  and R<sup>2</sup> of 0.698, in comparison to the random forest method that gave an RMSE of 7.904  $\mu\text{g m}^{-3}$  and R<sup>2</sup> = 0.619. It should be noted that their neural network was limited when estimating PM for cloudy areas and the night-time. A simple innovative way to estimate real-time pollution levels without the need for monitoring stations was proposed by Kow *et al* who used an image-based deep learning ResNet and VGG models that integrated a CNN and regression classification layer to create a relationship with PM<sub>2.5</sub> and PM<sub>10</sub> datasets collected at an air quality monitoring station in Taiwan, to estimate air pollution from cityscape photographs [110]. The CNN-RC was able to classify and estimate PM<sub>2.5</sub>, PM<sub>10</sub> and AQI values at the same time based on multiple inputs, and achieved a classification accuracy for PM<sub>2.5</sub>, PM<sub>10</sub> and AQI based on day and night images above 70%. Even when the images were less clear and there was presence of high levels of smoke, the CNN-RC still produced accurate estimates on pollutant concentrations.

#### 4.2.4. Forecasting AI

Deep learning CNNs have been used for 1 day to 7 day forecasting of pollen grains of tree, weed and grass pollen [111]. Meteorological data such as wind speed, total precipitation and temperature were used along with pollen data from archives in Houston, U.S.A. For comparison with daily observations, the neural network was able to predict with a relatively high index of agreement of 0.9, and the CNN showed promising ability for multiple day



predictions when compared to the conventional modelling approach. The predictions of seasonal trends in weed pollen within the area of study also correlated strongly with observations, but the model experienced a slight decrease in prediction accuracy as the forecast time was increased.

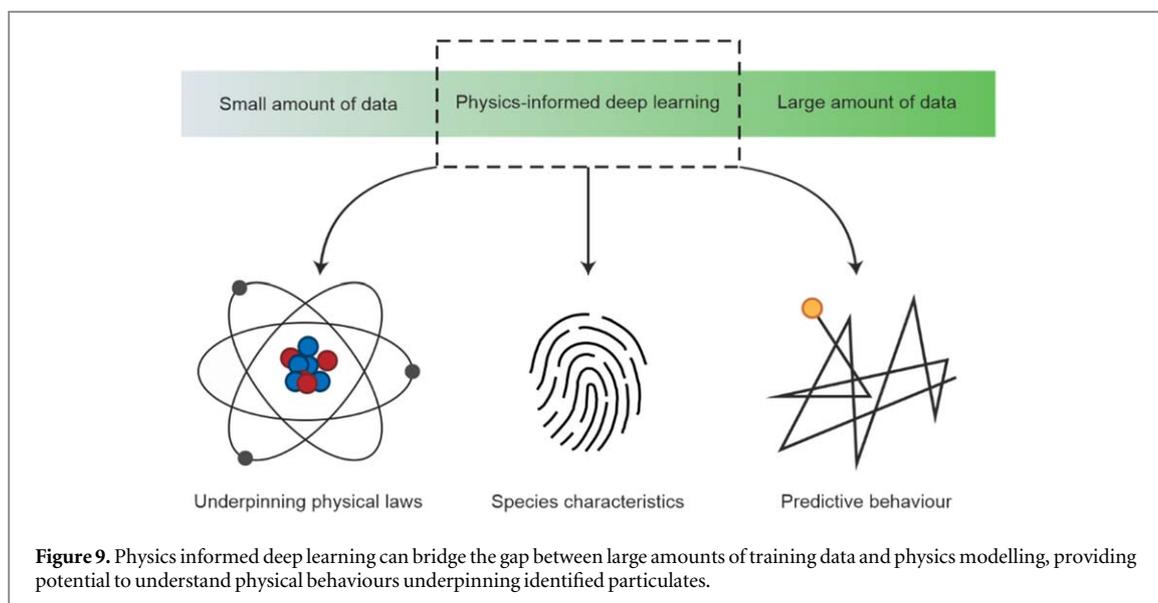
PM<sub>2.5</sub> concentrations have also been predicted by using both CNN and LSTM neural networks to achieve more accurate results than previous models, whereby the CNN was used to extract spatial features from a 2D matrix of data (meteorological and historical pollution data), which were then fed into an LSTM used to extract time series features for the input [112]. This CNN + LSTM model was compared with classical recurrent neural network (RNN), LSTM and back-propagation (BP) based methods. The RMSE of the CNN + LSTM was  $14.3 \mu\text{g m}^{-3}$ , the RNN was  $30.66 \mu\text{g m}^{-3}$  and the BP was  $22.37 \mu\text{g m}^{-3}$ . The lower RMSE of the CNN + LSTM method indicated that incorporating time-series feature information based on the correlation of spatial features could improve long-term sequence prediction accuracy of particulate matter concentrations.

Aggarwal and Toshniwal [113] also used LSTM-based deep learning neural network for urban air quality forecasting at a variety of locations in India for monthly time periods, and used a particle swarm optimisation of the hyperparameters to further improve the accuracy. The results of the LSTM-based neural network gave a RMSE score of  $14.17 \mu\text{g m}^{-3}$ , which outperformed existing models such as autoregressive integrated moving average (ARIMA) [114] and radial basis function neural network (RBFNN) [115], which had an RMSE score of  $24.44 \mu\text{g m}^{-3}$  and  $44.33 \mu\text{g m}^{-3}$ , respectively. Evaluating the LSTM-based model gave a R<sup>2</sup>-score ranging from 0.86 to 0.99 fifteen locations in India, indicating high accuracy.

## 5. Future perspectives

A lot of the progress appears to be image based, particularly in satellite imagery or optoelectronics, while prediction accuracy is based on the ability to scale neural networks with more data to achieve higher accuracy. This is most likely due to a combination of the following two observations. Firstly, image-based neural networks have been developed extremely rapidly in past years due to easy access to billions of labelled images available online. Secondly, since current methods of identification of concentration levels can involve imaging the particulates individually or as a mass (e.g., photographs of skylines), application of neural networks has focussed on these areas. Since the accuracy of neural networks tend to be dependent on the amount of training data [116], accuracies of image-based neural networks are likely to improve over time as more data can be acquired.

From the authors perspective, an ideal sensor (see figure 8) is one that can provide real-time, species specific information, at low cost to the consumer, with a small footprint, whether it be for satellite deployment or for ground-based sensing, with the data accessible and readable for a wide variety of people, and potentially updatable to allow for new neural network capabilities, (i.e., 3D imaging or tracking new tree mould species). Of course, such a device should be as green as possible and could also be powered without mains electricity to allow for portability, in the case of ground-based and in-home sensing.

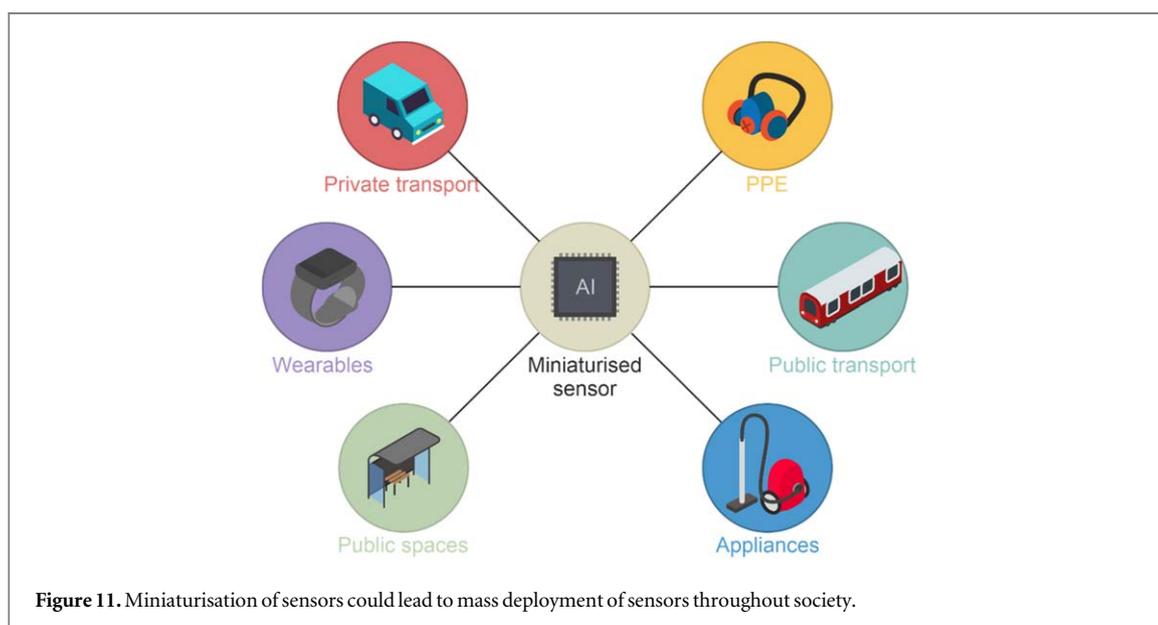
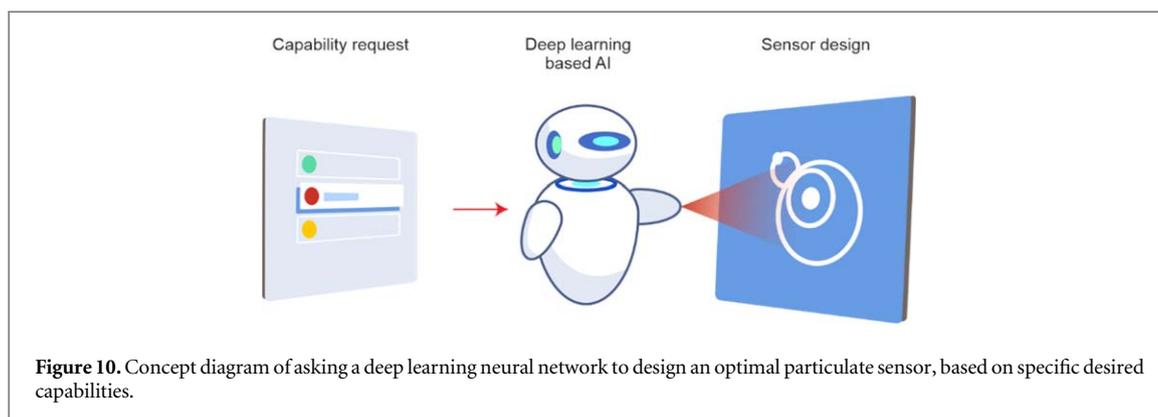


### 5.1. Physics informed models

Whilst future work could involve combining multiple neural networks for greater accuracy, as demonstrated by Luo *et al* [117], in which they combined convolutional neural networks, deep neural networks and integrated gradients for estimating ozone levels in the atmosphere using ground-truth data, physics-informed neural networks could provide scalable solutions to inverse problems with sparse data. In general, the greater the amount of data available to a neural network for training, the more likely it is to be able to develop an understanding of the physical laws present in the data, and hence achieve a higher prediction accuracy [118]. On the contrary, a purely physics-based model requires little data [119]. Hence, depending on the amount of data available to train neural networks, physics informed neural networks (PINNs) could be used to improve the capabilities and accuracy of the neural networks, leading to an understanding, in reference to this work, of the underpinning physical laws of the particulate sensing methodology, determination of particulate species characteristics and understanding particulate behaviour (see figure 9).

Recent work has shown that PINNs are effective and efficient for inverse problems, using incomplete datasets, and are scalable [120]. To date, magnetic resonant image (MRI) segmentation has been performed via combining MRI data with physics parameters in training of a neural network [121]. This is explained to outperform current segmentation deep learning techniques since the networks can produce segmentations that are robust to variations in the MRI physics, as the networks are learning the physics that contribute to the image generation. Other work includes integrating physics computer models into the deep learning training process, such as incorporating phase recovery physics computation into deep learning training to enable optimisation of the microscope illumination source pattern for Quantitative Phase Imaging (QPI) of stain-free and label-free microscopy [122]. Since some signal data from sensing can be difficult to interpret, such as light scatter from particulates, PINNs could support deep learning neural networks in the identification of particulates. PINNs could also aid in atmospheric modelling and forecasting of particulate levels and particulate transport.

Current methods of particulate monitoring use either data driven or physics-based modelling [123]. Since pollution forecasting involves a high-dimensional parameter space, with partial amounts of data, the ability of deep learning to integrate data and physics models could provide suitable architecture for highly accurate and efficient forecasting models. Indeed, preliminary work in this emerging field has been carried out using the outputs of the physical model as the inputs for the deep-learning model, for the purpose of predicting fine-mode fractions (FMF) of aerosols over land, and achieving an accuracy higher than current methods [124]. The authors found that, compared with more traditional physical-based and deep-learning-based FMF results, their PINN was more accurate at predicting FMF values over five land types (e.g., barren land, croplands, forests, grasslands and urban area). Kashinath *et al* [125] surveyed neural network models that incorporate physics and domain knowledge, to develop more generalised climate and weather forecasting weather predictions with greater consistency, improved data efficiency and reduced training times. Their survey found that PINNs had a reduced training time and greater data efficiency compared to previous models, gave greater physical consistency, and provided a better generalisation when forecasting weather and climate processes.



## 5.2. Aided sensor design

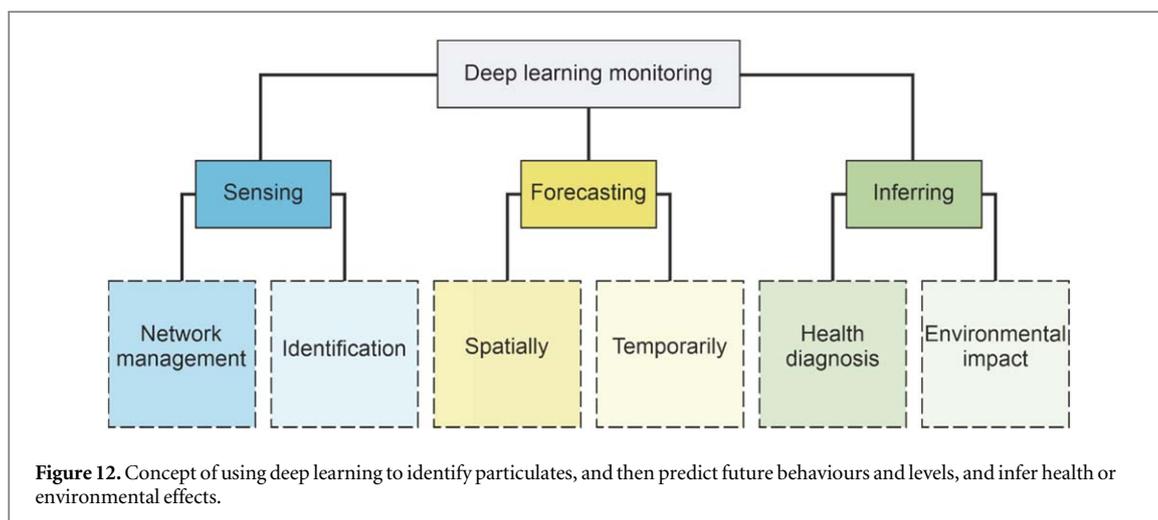
Future work could consist of using deep learning neural networks to design and optimise sensors. For example, deep neural networks have been used to optimise photonic crystal nanocavities [126], design and characterise plasmonic nanostructures [127], design chiral metamaterials [128] and enable the design of fibres for coherent beam combination [129]. As conceptualised in figure 10, in the future, one could ask an AI how to best design a sensor capable of low-cost sensing and identification of single particulate species.

## 5.3. Computer miniaturisation

Future deep learning in the field of particulate matter sensing will likely involve larger training sets for more accuracy [116]. This will necessitate more powerful GPUs capable of carrying out training with more data and higher resolution in relatively workable times [130]. The trends suggest GPUs are becoming more efficient with  $\sim 100$  GFLOPS per Watt currently available [76]. Also, for sensor footprint reduction, which is necessary for mass deployment and thus higher spatial resolution, increasingly capable microcomputers and smaller GPUs will also be likely necessary. Indeed, to date, CNNs using a Raspberry Pi [131] have been used for image recognition, such as for identifying animals in images recorded by a camera in the wild [132], and for classifying microbeads suspended in water [133]. As displayed in figure 11, one can envisage that, depending on the type of sensors developed, sensors could eventually be miniaturised through AI aided design and implemented in smart watches, vehicles, personal protective equipment (PPE), air purifiers and vacuum cleaners, as well as public areas such as bus shelters.

## 5.4. Monitoring

Beyond sensor design, deep reinforcement learning could aid in real-time monitoring, decision making and analysis, as described by Visez *et al* [134], as well as maintaining the sensor network, akin to magnetic control of tokamak plasmas as described by DeepMind [135]. In addition, deep learning could be used in the discovery of



relationships between pollution monitoring and health monitoring, such as linking pollution levels on a societal and individual level to specific health concerns. This could enable diagnosing an individual's allergy to a species type of pollen or fungal spore. For example, in the case of a patient developing symptoms of rhinitis, neural networks could be used to analyse the air pollution nearby and attempt to identify a correlation with such health events and airborne particle events. Figure 12 shows the potential different aspects that an all-encompassing deep learning-based monitoring system could be designed to achieve, from particulate detection, to forecasting and inferring illnesses and, more generally, environmental impact.

## 6. Conclusion

This review has discussed current techniques for airborne particulate matter pollution sensing and recent advances in deep learning that can aid in (1) improving the accuracy of current methods, (2) developing new methods of detection, imaging and characterisation and (3) enabling future pathways for increasing spatial resolution, temporal resolution and enable mass deployment. These include using deep learning to identify particulates with low-cost devices, to identify particulates more accurately using analysis techniques such as fluorescence spectroscopy, improving the capability of devices via image transformation, and improving the accuracy of particulate levels in forecasting.

Recent results in the identification of particulate and their concentration have confirmed that the application of deep learning enables considerably higher prediction accuracies for both spatial and temporal dimensions. Greater amounts of high quality data is key to mitigation, forecasting and understanding health effects, and to obtain more data, sensors with higher accuracy and greater spatial and temporal resolution should be developed. We anticipate that as neural networks become larger in size and are trained on greater amounts of data, further improvements to forecasting and particulate identification (which are multidimensional spatial-temporal problems) will be achieved. We conclude that deep learning has only just begun to impact particulate matter sensing, and indeed deep learning will be used to solve sensing challenges in domains not yet considered.

However, there is considerable opportunity for improvement, particularly in the domain of species level identification in real-time on low cost ground-based devices, for space-based sensing, and for global forecasting of particulate matter levels. The physical size of the sensing devices also needs to be reduced in size to enable larger distribution of devices, for enabling higher spatial resolution and for implementation in other products such as PPE and transportation. Furthermore, there is much work needed to be done for the real-time identification of airborne particulates that are smaller than PM<sub>2.5</sub>. This will perhaps require a combination of existing sensing methods with new technologies. One possibility here is the development of shorter wavelength sensors (extreme ultraviolet etc.) for enabling species level imaging or detection below the diffraction limit for visible wavelength light, or multispectral and multi-modal identification could be employed, where light, sound or electronic sensors are combined. To achieve this, as deep learning continues to become a more established tool, it is possible deep learning will be applied to the development of the design of the sensors themselves, for both abstract design of new types of sensing methods and instrumentation, and for miniaturisation of current techniques.

Finally, owing to the transportation of particulate matter around the globe, and the effect of different particulates on the human health, an important consideration here is to ensure particulate matter data are made freely available online. This will enable researchers to develop more accurate modelling, begin to move towards

the standardisation of data collection and processing, and allow public access to data for assisting in local and global policy design. In addition, the data will need to be correlated with health data to aid in health diagnosis and source identification.

## Acknowledgments

This work was funded by the Engineering and Physical Sciences Research Council (grant no. EP/T026197/1). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

## Conflicts of interest

The authors declare no conflict of interest.

## ORCID iDs

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

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

## References

- [1] Nriagu J O 1989 A global assessment of natural sources of atmospheric trace metals *Nature* **338** 47–9
- [2] Mukherjee A and Agrawal M 2017 World air particulate matter: sources, distribution and health effects *Environ. Chem. Lett.* **15** 283–309
- [3] Cha Y and Olofsson U 2018 Effective density of airborne particles in a railway tunnel from field measurements of mobility and aerodynamic size distributions *Aerosol Sci. Technol.* **52** 886–99
- [4] Zuurbier M, Hoek G, Oldenwening M, Lenters V, Meliefste K, van den Hazel P and Brunekreef B 2010 Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route *Environ. Health Perspect.* **118** 783–9
- [5] Cadle S H, Mulawa P A, Hunsanger E C, Nelson K, Ragazzi R A, Barrett R, Gallagher G L, Lawson D R, Knapp K T and Snow R 1999 Composition of light-duty motor vehicle exhaust particulate matter in the denver, colorado area *Environ. Sci. Technol.* **33** 2328–39
- [6] Thorpe A and Harrison R M 2008 Sources and properties of non-exhaust particulate matter from road traffic: a review *Sci. Total Environ.* **400** 270–82
- [7] Bond T C *et al* 2013 Bounding the role of black carbon in the climate system: a scientific assessment *Journal of Geophysical Research: Atmospheres* **118** 5380–552
- [8] Venkataraman C, Habib G, Eiguren-Fernandez A, Miguel A H and Friedlander S K 2005 Residential biofuels in south asia: carbonaceous aerosol emissions and climate impacts *Science* (1979) **307** 1454–6
- [9] Maenhaut W, Vermeylen R, Claeys M, Vercauteren J and Roekens E 2016 Sources of the PM10 aerosol in flanders, belgium, and re-assessment of the contribution from wood burning *Sci. Total Environ.* **562** 550–60
- [10] Mazzei F, D'Alessandro A, Lucarelli F, Nava S, Prati P, Valli G and Vecchi R 2008 Characterization of particulate matter sources in an urban environment *Sci. Total Environ.* **401** 81–9
- [11] Eugênio Toledo V *et al* 2007 Evaluation of levels, sources and distribution of toxic elements in PM10 in a suburban industrial region, Rio de Janeiro, Brazil *Environ. Monit. Assess.* **139** 49–59
- [12] Al-Neaimi Y, Gomes J and Lloyd O L 2001 Respiratory illnesses and ventilatory function among workers at a cement factory in a rapidly developing country *Occup Med (Chic Ill)* **51** 367–73
- [13] Perrino C, Canepari S, Catrambone M, Dalla Torre S, Rantica E and Sargolini T 2009 Influence of natural events on the concentration and composition of atmospheric particulate matter *Atmos. Environ.* **43** 4766–79
- [14] Hyde P and Mahalov A 2020 Contribution of bioaerosols to airborne particulate matter *J. Air Waste Manage. Assoc.* **70** 71–7
- [15] Dacunto P J, Cheng K C, Acevedo-Bolton V, Jiang R T, Klepeis N E, Repace J L, Ott W R and Hildemann L M 2013 Real-time particle monitor calibration factors and PM 2.5 emission factors for multiple indoor sources *Environ Sci Process Impacts* **15** 1511–9
- [16] Park J *et al* 2017 Spatial-temporal dispersion of aerosolized nanoparticles during the use of consumer spray products and estimates of inhalation exposure *Environ. Sci. Technol.* **51** 7624–38
- [17] Chakraborty R, Heydon J, Mayfield M and Mihaylova L 2020 Indoor air pollution from residential stoves: examining the flooding of particulate matter into homes during real-world use *Atmosphere* **11** 1326
- [18] Li Z, Wen Q and Zhang R 2017 Sources, health effects and control strategies of indoor fine particulate matter (PM2.5): a review *Sci. Total Environ.* **586** 610–22
- [19] Isiugo K, Jandarov R, Cox J, Ryan P, Newman N, Grinshpyn S A, Indugula R, Vesper S and Reponen T 2019 Indoor particulate matter and lung function in children *Sci. Total Environ.* **663** 408–17
- [20] McCormack M C, Breyse P N, Hansel N N, Matsui E C, Tonorez E S, Curtin-Brosnan J, Williams D L, Buckley T J, Eggleston P A and Diette G B 2008 Common household activities are associated with elevated particulate matter concentrations in bedrooms of inner-city Baltimore pre-school children *Environ. Res.* **106** 148–55

- [21] Sridharan S, Kumar M, Singh L, Bolan N S and Saha M 2021 Microplastics as an emerging source of particulate air pollution: a critical review *J. Hazard. Mater.* **418** 126245
- [22] Zhang L *et al* 2021 Indoor particulate matter in urban households: sources, pathways, characteristics, health effects, and exposure mitigation *Int J Environ Res Public Health* **18** 11055
- [23] 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
- [24] Ličbinský R, Frýbort A, Huzlík J, Adamec V, Effenberger K, Mikuška P, Vojtěšek M and Křůmal K 2010 Usage of scanning electron microscopy for particulate matter sources identification *Transactions on Transport Sciences* **3** 137–44
- [25] Raes F, van Dingenen R, Vignati E, Wilson J, Putaud J P, Seinfeld J H and Adams P 2000 Formation and cycling of aerosols in the global troposphere *Atmos. Environ.* **34** 4215–40
- [26] Visez N, Ivanovsky A, Roose A, Gosselin S, Sénéchal H, Poncet P and Choël M 2020 Atmospheric particulate matter adhesion onto pollen: a review *Aerobiologia (Bologna)* **36** 49–62
- [27] Parvez F, Lamancusa C and Wagstrom K 2017 Primary and secondary particulate matter intake fraction from different height emission sources *Atmos. Environ.* **165** 1–11
- [28] Spindler G, Grüner A, Müller K, Schlimper S and Herrmann H 2013 Long-term size-segregated particle (PM<sub>10</sub>, PM<sub>2.5</sub>, PM<sub>1</sub>) characterization study at Melpitz - Influence of air mass inflow, weather conditions and season *J. Atmos. Chem.* **70** 165–95
- [29] Vieno M *et al* 2016 The UK particulate matter air pollution episode of March–April 2014: more than Saharan dust *Environ. Res. Lett.* **11** 044004
- [30] Bathmanabhan S and Saragur Madanayak S N 2010 Analysis and interpretation of particulate matter—PM<sub>10</sub>, PM<sub>2.5</sub> and PM<sub>1</sub> emissions from the heterogeneous traffic near an urban roadway *Atmos Pollut Res* **1** 184–94
- [31] Clements A L, Fraser M P, Upadhyay N, Herckes P, Sundblom M, Lantz J and Solomon P A 2014 Chemical characterization of coarse particulate matter in the desert southwest—pinal county arizona, USA *Atmos Pollut Res* **5** 52–61
- [32] Munir S 2017 Analysing temporal trends in the ratios of PM<sub>2.5</sub>/PM<sub>10</sub> in the UK *Aerosol Air Qual. Res.* **17** 34–48
- [33] Anderson J O, Thundiyil J G and Stolbach A 2012 Clearing the air: a review of the effects of particulate matter air pollution on human health *J. Med. Toxicol.* **8** 166–75
- [34] Holgate S T 2017 Every breath we take: the lifelong impact of air pollution—a call for action *Clinical Medicine* **17** 8–12
- [35] Burnett R *et al* 2018 Global estimates of mortality associated with longterm exposure to outdoor fine particulate matter *Proc Natl Acad Sci USA* **115** 9592–7
- [36] Loxham M, Davies D E and Holgate S T 2019 The health effects of fine particulate air pollution *Brit. Med. J.* **367** 16609
- [37] Pope C A I, Ezzati M and Dockery D W 2009 Fine-particulate air pollution and life expectancy in the United States *The New England Journal of Medicine* **360** 376–86
- [38] Li L, Xing C, Zhou J, Niu L, Luo B, Song M, Niu J, Ruan Y, Sun X and Lei Y 2021 Airborne particulate matter (PM<sub>2.5</sub>) triggers ocular hypertension and glaucoma through pyroptosis *Part. Fibre Toxicol.* **18** 1–13
- [39] World Allergy Organization 2013 WAO White Book on Allergy 2013 Update ed R Pawankar, S T Holgate, W Canonica, R F Lockey and M S Blaiss (World Allergy Organization)
- [40] Zeleke Z K, Moen B E and Brätveit M 2010 Cement dust exposure and acute lung function: A cross shift study *BMC Pulm Med* **10** 19
- [41] Jenner L C, Rotchell J M, Bennett R T, Cowen M, Tentzeris V and Sadofsky L R 2022 Detection of microplastics in human lung tissue using  $\mu$ FTIR spectroscopy *Sci. Total Environ.* **831** 154907
- [42] Arden Pope C, Burnett R T, Turner M C, Cohen A, Krewski D, Jerrett M, Gapstur S M and Thun M J 2011 Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: Shape of the exposure-response relationships *Environ. Health Perspect.* **119** 1616–21
- [43] Rohr A C, Habre R, Godbold J, Moshier E, Schachter N, Kattan M, Grunin A, Nath A, Coull B and Koutrakis P 2014 Asthma exacerbation is associated with particulate matter source factors in children in New York City *Air Qual. Atmos. Health* **7** 239–50
- [44] VoPham T, Kim N J, Berry K, Mendoza J A, Kaufman J D and Ioannou G N 2022 PM<sub>2.5</sub> air pollution exposure and nonalcoholic fatty liver disease in the nationwide inpatient sample *Environ. Res.* **213** 113611
- [45] Taj T, Poulsen A H, Ketzel M, Geels C, Brandt J, Christensen J H, Puett R, Hvidtfeldt U A, Sørensen M and Raaschou-Nielsen O 2020 Long-term exposure to PM<sub>2.5</sub> and its constituents and risk of non-hodgkin lymphoma in Denmark: a population-based case-control study *Environ. Res.* **188** 109762
- [46] 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 USA* **113** 10797–801
- [47] Jung C R, Lin Y T and Hwang B F 2015 Ozone, particulate matter, and newly diagnosed alzheimer's disease: a population-based cohort study in Taiwan *Journal of Alzheimer's Disease* **44** 573–84
- [48] Oudin A, Segerström D, Adolfsson R and Forsberg B 2018 Association between air pollution from residential wood burning and dementia incidence in a longitudinal study in Northern Sweden *PLoS One* **13** e0198283
- [49] Lin H *et al* 2017 Ambient PM<sub>2.5</sub> and stroke: effect modifiers and population attributable risk in six low- and middle-income countries *Stroke* **48** 1191–7
- [50] Volk H E, Lurmann F, Penfold B, Hertz-Picciotto I and McConnell R 2013 Traffic-related air pollution, particulate matter, and autism *JAMA Psychiatry* **70** 71–7
- [51] von Schneidmesser E *et al* 2015 Chemistry and the linkages between air quality and climate change *Chem. Rev.* **115** 3856–97
- [52] Ramanathan V and Carmichael G 2008 Global and regional climate changes due to black carbon *Nat. Geosci.* **1** 221–7
- [53] Hadley O L and Kirchstetter T W 2012 Black-carbon reduction of snow albedo *Nat. Clim. Change* **2** 437–40
- [54] Rai P K 2016 Impacts of particulate matter pollution on plants: implications for environmental biomonitoring *Ecotoxicol. Environ. Saf.* **129** 120–36
- [55] Luo X, Bing H, Luo Z, Wang Y and Jin L 2019 Impacts of atmospheric particulate matter pollution on environmental biogeochemistry of trace metals in soil-plant system: a review *Environ. Pollut.* **255** 113138
- [56] Fourati R, Scopa A, ben Ahmed C, ben Abdallah F, Terzano R, Gattullo C E, Allegretta I, Galgano F, Caruso M C and Sofò A 2017 Leaf biochemical responses and fruit oil quality parameters in olive plants subjected to airborne metal pollution *Chemosphere* **168** 514–22
- [57] Grantz D A, Garner J H B and Johnson D W 2003 Ecological effects of particulate matter *Environ. Int.* **29** 213–39
- [58] Das S, Pal D and Sarkar A 2021 Particulate matter pollution and global agricultural productivity *Sustainable Agriculture Reviews* **50** 79–107
- [59] W. H. O. (WHO), 'Ambient (outdoor) air pollution,' ([https://who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health))

- [60] Ali M U, Liu G, Yousaf B, Ullah H, Abbas Q and Munir M A M 2018 A systematic review on global pollution status of particulate matter-associated potential toxic elements and health perspectives in urban environment *Environ. Geochem. Health* **41** 1131–62
- [61] 'Federal register :: national ambient air quality standards for particulate matter,' (<https://federalregister.gov/documents/2006/10/17/06-8477/national-ambient-air-quality-standards-for-particulate-matter>)
- [62] Patashnick H and Hemenway C L 1969 Oscillating fiber microbalance *Rev. Sci. Instrum.* **40** 1008–11
- [63] Kulkarni P, Baron P A and Willeke K 2011 *Aerosol measurement: principles, techniques, and applications* (John Wiley & Sons, Inc.) 3rd edn (<https://doi.org/10.1002/9781118001684>)
- [64] Schweizer D, Cisneros R and Shaw G 2016 A comparative analysis of temporary and permanent beta attenuation monitors: the importance of understanding data and equipment limitations when creating PM<sub>2.5</sub> air quality health advisories *Atmos Pollut Res.* **7** 865–75
- [65] Gupta P and Christopher S A 2009 Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach *Journal of Geophysical Research: Atmospheres* **114** 14205
- [66] Engel-Cox J A, Hoff R M, Rogers R, Dimmick F, Rush A C, Szykman J J, Al-Saadi J, Chu D A and Zell E R 2006 Integrating lidar and satellite optical depth with ambient monitoring for 3-dimensional particulate characterization *Atmos. Environ.* **40** 8056–67
- [67] Kaliszewski M, Włodarski M, Młyniczak J and Kopczyński K 2020 Comparison of low-cost particulate matter sensors for indoor air monitoring during covid-19 lockdown *Sensors (Basel)* **20** 1–17
- [68] Giordano M R, Malings C, Pandis S N, Presto A A, McNeill V F, Westervelt D M, Beekmann M and Subramanian R 2021 From low-cost sensors to high-quality data: a summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors *J. Aerosol Sci.* **158** 105833
- [69] Neumeister-Kemp H G, Maxwell A A, Kemp P C, Dell B, St G E and Hardy J 2004 An advanced slit-type volumetric spore trap for monitoring bioaerosols; new methods for identifying fungal spores *Australasian Plant Pathology* **33** 393–400
- [70] Sterling M, Rogers C and Levetin E 1999 An evaluation of two methods used for microscopic analysis of airborne fungal spore concentrations from the burkard spore trap *Aerobiologia (Bologna)* **15** 9–18
- [71] Ghosh N, Whiteside M, Ridner C, Celik Y, Saadeh C, Nabarun Ghosh A and Bennert J 2010 Characterization of aeroallergen of Texas panhandle using scanning and fluorescence microscopy *Proceedings Volume 7729 Scanning Microscopy 2010* 7729, 401–9
- [72] Hutton B M and Williams D E 2000 Assessment of x-ray photoelectron spectroscopy for analysis of particulate pollutants in urban air *Analyst* **125** 1703–6
- [73] Schmidhuber J 2015 Deep learning in neural networks: an overview *Neural Netw.* **61** 85–117
- [74] McCulloch W S and Pitts W 1943 A logical calculus of the ideas immanent in nervous activity *The Bulletin of Mathematical Biophysics* **5** 115–33
- [75] Dally W J, Keckler S W and Kirk D B 2021 Evolution of the graphics processing unit (GPU) *IEEE Micro* **41** 42–51
- [76] Sun Y, Agostini N B, Dong S and Kaeli D 2019 Summarizing CPU and GPU design trends with product data (<https://doi.org/10.48550/arXiv.1911.11313>)
- [77] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in Neural Information Processing Systems* **25** 1097–105
- [78] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets,' in *Advances in Neural Information Processing Systems* **27** 2672–80
- [79] Isola P, Zhu J-Y, Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 5967–76
- [80] Ruff Z J, Lesmeister D B, Duchac L S, Padmaraju B K and Sullivan C M 2020 Automated identification of avian vocalizations with deep convolutional neural networks *Remote Sens Ecol Conserv* **6** 79–92
- [81] 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
- [82] Bergler C, Schröter H, Cheng R X, Barth V, Weber M, Nöth E, Hofer H and Maier A 2019 ORCA-SPOT: an automatic killer whale sound detection toolkit using deep learning *Sci. Rep.* **9** 1–17
- [83] Daood A, Ribeiro E and Bush M 2016 Pollen grain recognition using deep learning *Int. Symp. on Visual Computing* **10072** 321–30
- [84] Cortes C, Vapnik V and Saitta L 1995 Support-vector networks *Mach. Learn.* **20** 273–97
- [85] Sevillano V, Holt K and Aznarte J L 2020 Precise automatic classification of 46 different pollen types with convolutional neural networks *PLoS One* **15** 1–15
- [86] Gonçalves A B, Souza J S, da Silva G G, Cereda M P, Pott A, Naka M H and Pistori H 2016 Feature extraction and machine learning for the classification of brazilian savannah pollen grains *PLoS One* **11** e0157044
- [87] Oteros J, Weber A, Kutzora S, Rojo J, Heinze S, Herr C, Gebauer R, Schmidt-Weber C B and Buters J T M 2020 An operational robotic pollen monitoring network based on automatic image recognition *Environ. Res.* **191** 110031
- [88] Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z 2016 Rethinking the inception architecture for computer vision *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2818–26
- [89] Schaefer J, Milling M, Schuller B W, Bauer B, Brunner J O, Traidl-Hoffmann C and Damialis A 2021 Towards automatic airborne pollen monitoring: from commercial devices to operational by mitigating class-imbalance in a deep learning approach *Sci. Total Environ.* **796** 148932
- [90] Wu Y *et al* 2018 Label-free bioaerosol sensing using mobile microscopy and deep learning *ACS Photonics* **5** 4617–27
- [91] Yin C, Cheng X, Liu X and Zhao M 2020 Identification and classification of atmospheric particles based on SEM images using convolutional neural network with attention mechanism *Complexity* **2020** 1–13
- [92] D'Orazio M, Reale R, de Ninno A, Brighetti M A, Mencattini D A, Businaro L, Martinelli E, Bisegna P, Travaglini A and Caselli F 2022 Electro-optical classification of pollen grains via microfluidics and machine learning *IEEE Trans. Biomed. Eng.* **69** 921–31
- [93] Kim J, Go T and Lee S J 2021 Volumetric monitoring of airborne particulate matter concentration using smartphone-based digital holographic microscopy and deep learning *J. Hazard. Mater.* **418** 126351
- [94] 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
- [95] Gorodkin J 2004 Comparing two K-category assignments by a K-category correlation coefficient *Comput. Biol. Chem.* **28** 367–74
- [96] 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
- [97] Grant-Jacob J A, Praeger M, Loxham M, Eason R W and Mills B 2020 Lensless imaging of pollen grains at three-wavelengths using deep learning *Environ Res Commun* **2** 75005

- [98] Grant-Jacob J A, Mackay B S, Baker J A G, Xie Y, Heath D J, Loxham M, Eason R W and Mills B 2019 A neural lens for super-resolution biological imaging *J. Phys. Commun.* **3** 065004
- [99] Grant-Jacob J A, Praeger M, Eason R W and Mills B 2022 Generating images of hydrated pollen grains using deep learning *IOP SciNotes* **3** 024001
- [100] Daunys G, Šukienė L, Vaitkevičius L, Valiulis G, Sofiev M and Šaulienė I 2021 Clustering approach for the analysis of the fluorescent bioaerosol collected by an automatic detector *PLoS One* **16** e0247284
- [101] Tešendić D, Boberić Krstićev D, Matavulj P, Brdar S, Panić M, Minić V and Šikoparija B 2022 RealForAll: real-time system for automatic detection of airborne pollen *Enterprise Information Systems* **16** 1793391
- [102] Rutherford J W, Dawson-Elli N, Manicone A M, Korshin G V, Novosselov I V, Seto E and Posner J D 2020 Excitation emission matrix fluorescence spectroscopy for combustion generated particulate matter source identification *Atmos. Environ.* **220** 117065
- [103] Hasti V R and Shin D 2022 Denoising and fuel spray droplet detection from light-scattered images using deep learning *Energy and AI* **7** 100130
- [104] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–8
- [105] Imani M 2021 Particulate matter (PM<sub>2.5</sub> and PM<sub>10</sub>) generation map using MODIS Level-1 satellite images and deep neural network *J. Environ. Manage.* **281** 111888
- [106] Li T, Shen H, Yuan Q, Zhang X and Zhang L 2017 Estimating ground-level PM<sub>2.5</sub> by fusing satellite and station observations: a geointelligent deep learning approach *Geophys. Res. Lett.* **44** 985–11
- [107] Sun Y, Zeng Q, Geng B, Lin X, Sude B and Chen L 2019 Deep learning architecture for estimating hourly ground-level PM<sub>2.5</sub> using satellite remote sensing *IEEE Geosci. Remote Sens. Lett.* **16** 1343–7
- [108] Yan X, Zang Z, Luo N, Jiang Y and Li Z 2020 New interpretable deep learning model to monitor real-time PM<sub>2.5</sub> concentrations from satellite data *Environ Int* **144** 106060
- [109] Lee C, Lee K, Kim S, Yu J, Jeong S and Yeom J 2021 Hourly ground-level PM<sub>2.5</sub> estimation using geostationary satellite and reanalysis data via deep learning *Remote Sensing* **13** 2121
- [110] Kow P Y, Hsia I W, Chang L C and Chang F J 2022 Real-time image-based air quality estimation by deep learning neural networks *J. Environ. Manage.* **307** 114560
- [111] Lops Y, Choi Y, Eslami E and Sayeed A 2020 Real-time 7 day forecast of pollen counts using a deep convolutional neural network *Neural Comput Appl* **32** 11827–36
- [112] Qin D, Yu J, Zou G, Yong R, Zhao Q and Zhang B 2019 A novel combined prediction scheme based on CNN and LSTM for urban PM<sub>2.5</sub> concentration *IEEE Access* **7** 20050–9
- [113] Aggarwal A and Toshniwal D 2021 A hybrid deep learning framework for urban air quality forecasting *J. Clean. Prod.* **329** 129660
- [114] Diaz-Robles L A, Ortega J C, Fu J S, Reed G D, Chow J C, Watson J G and Moncada-Herrera J A 2008 A hybrid arima and artificial neural networks model to forecast particulate matter in urban areas: The case of temuco, Chile *Atmos. Environ.* **42** 8331–40
- [115] Song C and Fu X 2020 Research on different weight combination in air quality forecasting models *J. Clean. Prod.* **261** 121169
- [116] Soekhoe D, van der Putten P and Plaata A 2016 On the impact of data set size in transfer learning using deep neural networks *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **9897** 50–60
- [117] Luo N, Zang Z, Yin C, Liu M, Jiang Y, Zuo C, Zhao W, Shi W and Yan X 2022 Explainable and spatial dependence deep learning model for satellite-based O<sub>3</sub> monitoring in China *Atmos. Environ.* **290** 119370
- [118] Mills B, Heath D J, Grant-Jacob J A and Eason R W 2018 Predictive capabilities for laser machining via a neural network *Opt. Express* **26** 17245–53
- [119] Karniadakis G E, Kevrekidis I G, Lu L, Perdikaris P, Wang S and Yang L 2021 Physics-informed machine learning *Nature Reviews Physics* **3** 422–40
- [120] Raissi M, Perdikaris P and Karniadakis G E 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations *J. Comput. Phys.* **378** 686–707
- [121] Borges P, Sudre C, Varsavsky T, Thomas D, Drobnjak I, Ourselin S and Cardoso M J 2019 Physics-informed brain MRI segmentation *Simulation and Synthesis in Medical Imaging* (Springer International Publishing) 11827, 100–9
- [122] Kellman M, Bostan E, Repina N and Waller L 2019 Physics-based learned design: optimized coded-illumination for quantitative phase imaging *IEEE Trans Comput Imaging* **5** 344–53
- [123] Fablet R, Amar M M, Febvre Q, Beauchamp M and Chapron B 2021 End-to-end physics-informed representation learning for satellite ocean remote sensing data: applications to satellite altimetry and sea surface currents *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **5** 295–302
- [124] Yan X *et al* 2022 A global land aerosol fine-mode fraction dataset (2001–2020) retrieved from MODIS using hybrid physical and deep learning approaches *Earth Syst Sci Data* **14** 1193–213
- [125] Kashinath K *et al* 2021 Physics-informed machine learning: case studies for weather and climate modelling *Philos Trans A Math Phys Eng Sci* **379**
- [126] Noda S and Asano T 2018 Optimization of photonic crystal nanocavities based on deep learning *Opt. Express* **26** 32704–17
- [127] Malkiel I, Mrejen M, Nagler A, Arieli U, Wolf L and Suchowski H 2018 Plasmonic nanostructure design and characterization via deep learning *Light: Science & Applications* **7** 1–8
- [128] Ma W, Cheng F and Liu Y 2018 Deep-learning-enabled on-demand design of chiral metamaterials *ACS Nano* **12** 6326–34
- [129] Mills B, Grant-Jacob J A, Praeger M, Eason R W, Nilsson J and Zervas M N 2021 Single Step Phase Optimisation for Coherent Beam Combination Using Deep Learning *Scientific Reports* **12** 5188
- [130] Hestness J, Ardalani N and Diamos G 2019 Beyond human-level accuracy *Proceedings of the 24th Symposium on Principles and Practice of Parallel Programming* (New York: ACM) 1–14
- [131] Suzen A A, Duman B and Sen B 2020 Benchmark analysis of jetson TX2, jetson nano and raspberry PI using deep-CNN *HORA 2020 – 2nd Int. Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proc.* 1–5
- [132] Curtin B H and Matthews S J 2019 Deep learning for inexpensive image classification of wildlife on the raspberry Pi *2019 IEEE 10th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2019* 0082–7
- [133] 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** 35001
- [134] Skordilis E and Moghaddass R 2020 A deep reinforcement learning approach for real-time sensor-driven decision making and predictive analytics *Comput. Ind. Eng.* **147** 106600
- [135] Degraeve J *et al* 2022 Magnetic control of tokamak plasmas through deep reinforcement learning *Nature* **602** 414–9