Retrieval of Multiple Atmospheric Environmental Parameters From Images With Deep Learning

SU, Peifeng

2022-02-04

SU , P , Liu , Y , Tarkoma , S , Rebeiro-Hargrave , A , Petäjä , T , Kulmala , M & Pellikka , P 2022 , ' Retrieval of Multiple Atmospheric Environmental Parameters From Images With Deep Learning ' , IEEE Geoscience and Remote Sensing Letters , vol. 19 , 1005005 . https://doi.org/10.1109/LGRS.2

http://hdl.handle.net/10138/354290
https://doi.org/10.1109/LGRS.2022.3149045

acceptedVersion

Retrieval of Multiple Atmospheric Environmental Parameters from Images with Deep Learning

Peifeng Su, Yongchun Liu, Sasu Tarkoma, Andrew Rebeiro-Hargrave, Tuukka Petäjä, Markku Kulmala, and Petri Pellikka,

***Abstract*—Retrieving atmospheric environmental parameters such as atmospheric horizontal visibility and mass concentration of aerosol particles with diameter of 2.5 or 10 micrometers or less ($PM_{2.5}$, PM_{10}, respectively) from digital images provides new tools for horizontal environmental monitoring. In this study, we propose a new end-to-end convolutional neural network for the retrieval of multiple atmospheric environmental parameters (RMEP) from images. In contrast to other retrieval models, RMEP can retrieve a suite of atmospheric environmental parameters including atmospheric horizontal visibility, relative humidity, ambient temperature, $PM_{2.5}$, and PM_{10} simultaneously from a single image. Experimental results demonstrate that: (1) it is possible to simultaneously retrieve multiple atmospheric environmental parameters; (2) spatial and spectral resolutions of images are not the key factors for the retrieval on the horizontal scale; (3) RMEP achieves the best overall retrieval performance compared with several classic convolutional neural networks such as AlexNet, ResNet-50, and DenseNet-121, and the results are based on experiments on images extracted from webcams located in different continents (test R^2 values are 0.63, 0.72, and 0.82 for atmospheric horizontal visibility, relative humidity, and ambient temperature, respectively). Experimental results show the potential of utilizing webcams to help monitor the environment. Code and more results are available at https://github.com/cvvsu/RMEP.**

***Index Terms*—Environmental monitoring, atmospheric pollutant, meteorological parameters, convolutional neural network, deep learning, image processing.**

I. Introduction

MASS concentration of aerosol particles with diameter of 2.5 or 10 micrometers or less ($PM_{2.5}$, PM_{10}, respectively), aerosol number size distribution, aerosol chemical composition, (ambient) temperature, and relative humidity (RH) are key parameters governing the transmission of light through the urban troposphere and therefore governing (atmospheric horizontal) visibility. The aerosol mass concentration and visibility are tightly connected to air quality and particularly its regional variability in the urban landscape. To understand climate change and processes influencing air pollution, we need to make reliable observations about atmospheric pollutants and meteorological parameters. For simplicity, we term visibility, RH, temperature, $PM_{2.5}$, and PM_{10} as atmospheric environmental parameters in this study.

Peifeng Su, Sasu Tarkoma, Andrew Rebeiro-Hargrave, Tuukka Petäjä, Markku Kulmala, and Petri Pellikka are with University of Helsinki, FI-00014 Helsinki, Finland. e-mail: peifeng.su@helsinki.fi.

Yongchun Liu is with Beijing University of Chemical Technology, Beijing 100029, China.

Manuscript received September 15, 2021; revised December 14, 2021.

Ground-based comprehensive in-situ observations at fixed measurement locations are expensive and require well-trained operators for daily operations and maintenance. The observations are also tightly descriptive of the specific observation location. Particularly in the highly variable urban environments, the data are not representative for the whole city [1], [2]. For monitoring air quality, a network of observation stations is typically deployed to capture the spatio-temporal variability of the air pollution in the region. Another way forward to upscale the representativeness of the measurements is to utilize ground-based remote sensing such as Lidars [3], which are typically resolving the aerosol concentrations in the vertical dimension.

Analysis of digital images provides possible tools to capture air pollution on the horizontal scale [4]–[6] since digital images are influenced by atmospheric environmental parameters in many ways. More precisely, images captured under different environmental conditions can objectively record these conditions, thus offering the possibility to retrieve atmospheric environmental parameters. For example, images captured under hazy and foggy conditions manifest poor visibility, reduced contrast, fainted surfaces, and color shift [7].

To retrieve atmospheric environmental parameters from images, environment-related image features should be extracted from images, and then a regression or classification model can be built between these image features and reference values of atmospheric environmental parameters. For instance, image features such as dark channel and spatial contrast are related to visibility [8], [9] and $PM_{2.5}$ [4], and local pixel intensities are related to temperature [10]. Deep learning methods such as convolutional neural network (CNN) and recurrent neural network (RNN) can directly build relationships between images and reference values of atmospheric environmental parameters without extracting image features manually. CNN usually deals with single image [5], [11]–[14], while RNN is used to find relationships between sequential images [5]. CNN and RNN can also be used together for the retrieval [15], [16]. It is also possible to extract environment-related image features manually and then send these image features to a CNN, and the CNN can be regard as a regressor or a classifier in this case [6], [16], [17]. However, whether the spatial and spectral resolutions of images affect the retrieval results and whether it is possible to retrieve as many atmospheric environmental parameters as possible from images simultaneously are still unclear.

In this letter, we propose a new convolutional neural network which can estimate multiple atmospheric environmental

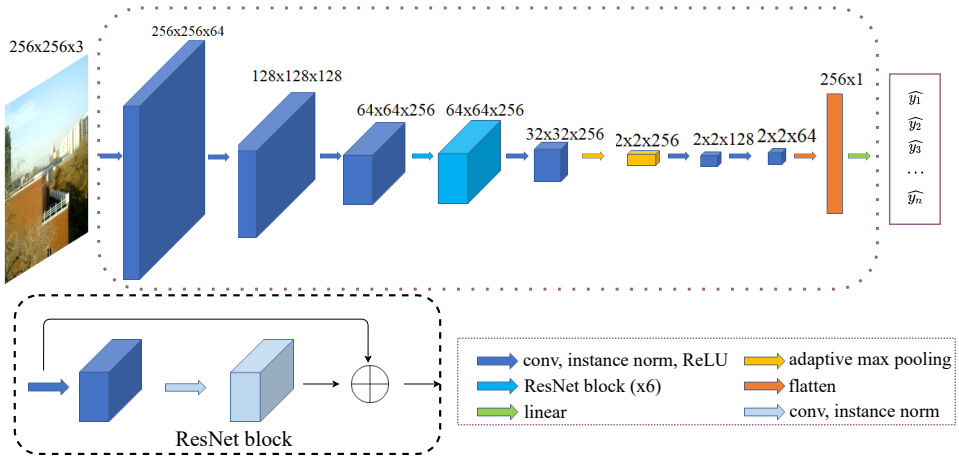

Fig. 1. Illustration of the RMEP model, assuming the spatial size of the input RGB image is 256×256 pixels. ResNet block is the basic component of ResNet [18]. *Instance norm* means the instance normalization [19].

parameters including visibility, RH, temperature, $PM_{2.5}$, and PM_{10} simultaneously. The main findings and contributions of this letter are:

1) it is possible to retrieve a suite of atmospheric environmental parameters from images simultaneously in an end-to-end manner;
2) on the horizontal scale, the spatial and spectral resolutions of images are not the key factors for the retrieval;
3) the proposed RMEP model achieves the best overall performance compared with several classic CNNs such as AlexNet [20], ResNet-50 (50-layer ResNet) [18], and DenseNet-121 (121-layer DenseNet) [21], in terms of the retrieval accuracy.

II. Methodology

A. Images and Atmospheric Environmental Parameters

In the computer vision community, the atmospheric scattering model is widely used for the removal of haze and fog [22], [23]:

$$\begin{aligned} \mathbf{I}(\mathbf{x}) &= \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})), \\ t(\mathbf{x}) &= e^{-\beta d(\mathbf{x})}, \end{aligned} \tag{1}$$

where $\mathbf{I}$ is the observed image, $\mathbf{J}$ is the scene radiance without attenuation, $\mathbf{A}$ is the global atmospheric light, t is the medium transmission, β is the scattering coefficient of the atmosphere, and d is the scene depth. Koschmieder's law is used for daytime visibility estimation [24]

$$V = -\frac{\ln(0.05)}{\beta} \doteq \frac{3}{\beta}, \tag{2}$$

where V is visibility.

According to Eq. (2), if the attenuation coefficient β could be obtained from an observed image, we can retrieve the visibility from the image. According to [25], visibility is affected by $PM_{2.5}$ and RH. Therefore, it is possible to retrieve visibility, RH, temperature, $PM_{2.5}$, and PM_{10} simultaneously. However, it is clear that the relationships between images and the related atmospheric environmental parameters are complex. Fortunately, deep learning methods can retrieve atmospheric environmental parameters with a lack of sound physical models due to their capacity of approximating complex relationships [26], [27]. Thus, we propose an end-to-end CNN for retrieving a suite of atmospheric environmental parameters without knowing the explicit relationship between them.

B. The RMEP Model and Smooth L_1 Loss

CycleGan has been presenting excellent performance in changing the weather conditions for images, such as image dehazing and transferring seasonal appearances of images [28], [29]. Specifically, CycleGan can convert a hazy image into a haze-free one and a haze-free image into a hazy one, meaning that CycleGan can automatically extract the visibility-related image features. Similarly, CycleGan can also extract temperature-related image features since it can transfer the seasonal appearances of images. Environmental-related features are important for the retrieval of atmospheric environmental parameters. Therefore, we crop the first few layers from the generator of CycleGan and add some new layers to build a model between images and atmospheric environmental parameters. We term the proposed model as RMEP for simplicity (Fig. 1).

Smooth L_1 loss l_{smooth} [30] is applied as the loss function

$$l_{smooth} = \begin{cases} 0.5 * (y - \hat{y})^2, & |y - \hat{y}| < 1 \\ |y - \hat{y}| - 0.5, & otherwise \end{cases} \tag{3}$$

where y is the reference value and $\hat{y}$ is the retrieved value.

III. Results and Discussion

A. Datasets

We use a portable hyperspectral device Specim IQ, which can record a hyperspectral and a related RGB image simultaneously, to collect paired hyperspectral and RGB images. A captured hyperspectral image is with the shape of 512×512 pixels and 204 channels, while a paired RGB image is with the shape of 645×645 pixels. The wavelength range of a hyperspectral image is 400-1000 nm.

TABLE I
DETAILED INFORMATION ABOUT IMAGE SIZES

Datasets	#Channel		Spatial size (pixels)		
	Original	Resampled	Original	Resampled	Training
RGB_645	3	3	645×645	645×645	640×640
RGB_512	3	3	645×645	512×512	510×510
RGB_256	3	3	645×645	256×256	254×254
RGB_128	3	3	645×645	128×128	126×126
HS3	204	3	512×512	128×128	126×126
HS30	204	30	512×512	128×128	126×126
HS51	204	51	512×512	128×128	126×126
HS117	204	117	512×512	128×128	126×126
HS204	204	204	512×512	128×128	126×126
SSF	3	3	⋆	160×120	118×118

⋆Original images in SSF dataset are with several different sizes.

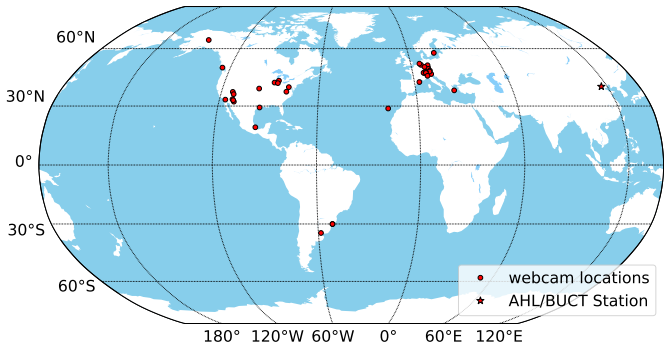

Fig. 2. Example images and the locations of the webcams and the AHL/BUCT Station.

Images are captured in the Aerosol and Haze Laboratory/Beijing University of Chemical Technology Station (AHL/BUCT Station) (Fig. 2) during the years 2018-2020, and the station provides the reference values for the above-mentioned atmospheric environmental parameters. After filtering out images without reference values, 751 paired images are obtained and termed as the RGBHSI dataset. We resample the RGB images to different spatial resolutions (512×512, 256×256, and 128×128 pixels, respectively) to check whether the spatial resolution will affect the retrieval results. Similarly, to verify whether the spectral resolution will affect the retrieval accuracy, we select three channels (449.35 nm, 548.55 nm, and 598.60 nm), the first 117 channels that are in the visible spectrum region (380-740 nm), and all the 204 channels from each hyperspectral image to construct the HS3, HS117, and HS204 datasets, respectively. Furthermore, we select channel 1, 5, 9, 13, 17, ..., 113, 117 from HS117 and channel 1, 5, 9, 13, 17, ..., 200, 204 from HS204 to construct the HS30 and HS51 datasets, respectively. Detailed spatial and channel information of images can be seen in TABLE I.

To validate the generality of the RMEP model, we use the SkyFinder dataset, which contains roughly 90,000 webcam images collected in different continents (Fig. 2). In [5], images in the SkyFinder dataset without reference values are dropped, and 35,417 images are left. We further drop images captured during nighttime and images with reference visibility values that are greater than 20,000 m. There are 24,328 images left, which construct the Sub-SkyFinder (SSF) dataset. The images in the SSF dataset are captured during the years 2011-2014. The reference values of the SkyFinder and the SSF datasets are from the weather stations that are nearby the locations where the webcams installed. Only reference values for visibility, RH, and temperature are available for the SSF dataset.

B. Experimental Setup

We run the experiments on the Finnish CSC Puhti Server, using one GPU (NVIDIA V100), and 16 CPUs (Xeon Gold 6230 @ 2.1 GHz). The learning rate is 2×10^{-4} for the first 50 epochs and then linearly decay to zero for the rest 150 epochs. Batch size is 16. Since RMEP is a CNN model, in general it cannot keep the sequential information, which can be extracted from sequential datasets by sequence models such as RNNs. Thus, we randomly split the datasets into training, validation, and test sets. Specifically, for each dataset, 20% of images are randomly selected as the test set, and out of the rest 80% images, 20% are of them randomly selected as the validation set, and the remaining images construct the training set. To avoid over-fitting, the resampled images are randomly cropped and horizontally flipped during training.

TABLE II
RETRIEVAL RESULTS ON THE TEST SETS IN RGBHSI

Datasets	Visibility		RH		Temperature		$PM_{2.5}$		PM_{10}	
	R^2	MAE (m)	R^2	MAE (%)	R^2	MAE (°C)	R^2	MAE ($\mu g/m^3$)	R^2	MAE ($\mu g/m^3$)
HS3	0.97	566.99	0.92	2.46	0.98	0.95	0.94	5.48	0.94	7.47
HS30	0.97	500.54	0.93	2.39	0.98	0.78	0.94	5.30	0.93	7.75
HS51	0.97	522.35	0.93	2.24	0.98	0.89	0.95	4.97	0.95	6.76
HS117	0.97	554.94	0.93	2.42	0.98	0.94	0.94	5.61	0.92	8.56
HS204	0.96	550.74	0.93	2.21	0.98	0.90	0.93	5.81	0.94	7.63
RGB_645	0.99	484.00	0.95	2.29	0.98	0.91	0.91	7.05	0.93	8.09
RGB_512	**0.99**	**344.29**	**0.96**	**1.74**	0.98	0.72	**0.95**	**4.83**	0.96	5.93
RGB_256	0.98	397.70	0.96	1.87	0.99	0.64	0.94	5.68	0.96	6.90
RGB_128	0.98	376.23	0.96	1.92	**0.99**	**0.62**	0.95	5.22	**0.97**	**6.07**

To evaluate the performance of the RMEP model, we utilize the coefficient of determinations (R^2) and the mean absolute error (MAE) to compare the retrieved and reference values. Please refer to our code to see more results.

C. Retrieval Results

The retrieval results on the test set of each dataset are shown in TABLE II. On the RGBHSI dataset, all the R^2 values are greater than 0.91, and the MAE values are less than 566.99 m, 2.46%, 0.95 °C, 7.05 $\mu g/m^3$, and 8.56 $\mu g/m^3$ for visibility, RH, temperature, $PM_{2.5}$, and PM_{10}, respectively, indicating that it is possible to retrieve multiple atmospheric environmental parameters from images simultaneously. Though atmospheric environmental parameters are retrieved at the same time, the retrieval accuracies of visibility and temperature are better than the retrieval accuracies of RH, $PM_{2.5}$, and PM_{10}.

Interestingly, compared with the hyperspectral datasets (HS3, HS30, HS1, HS117, and HS204), the RGB datasets (RGB_645, RGB_512, RGB_256, and RGB_128) have similar retrieval results (TABLE II). In addition to the spectral

resolution, different spatial resolutions offer similar retrieval results (TABLE II).

On a larger spatial scale, it is still possible to retrieve the atmospheric environmental parameters (TABLE III and Fig. 3), which further validate the effectiveness of the RMEP model. The retrieval accuracy of temperature is better than that of RH, which agrees with the retrieval results on the RGBHSI dataset. However, the retrieval accuracy of visibility is worse than that of RH, which is different from the retrieval results on the RGBHSI dataset.

We use the AlexNet, ResNet-50, and DenseNet-121 to further demonstrate the performance of the proposed RMEP model (TABLE III). The RMEP model achieves the best overall performance compared with AlexNet, ResNet-50, and DenseNet-121.

D. Ablation Study

On the SSF dataset, different pooling layers and numbers of ResNet blocks are used for ablation study. No matter whether the adaptive max polling (Fig. 1) or adaptive average pooling is utilized, their retrieval accuracies are similar (TABLE IV). In addition, different numbers of the ResNet blocks (Fig. 1) have similar retrieval results, and there is a balance among the retrieval accuracies of different atmospheric environmental parameters as the number of ResNet blocks grows (TABLE IV). Thus, it appears that utilizing six ResNet blocks in the RMEP model is a superior option.

E. Discussions

According to the above-mentioned experimental results, it is possible to retrieve multiple atmospheric environmental parameters including visibility, RH, temperature, $PM_{2.5}$, and PM_{10} simultaneously from images. Experimental results demonstrate that the spectral and special resolutions are not the key factors for the retrieval on the horizontal scale.

Visibility, RH, temperature, $PM_{2.5}$, and PM_{10} can be retrieved at the same time since these atmospheric environmental parameters can influence each other. For example, when the PM_{10} values are high, then generally it means that the visibility value is low. Visibility and temperature are easier to retrieve than RH, $PM_{2.5}$, and PM_{10} (TABLE II). On the SSF dataset, the retrieval accuracy of visibility is lower than that of RH and temperature, and one possible reason is that some weather stations are unable to obtain precise visibility values when the values exceed specific distances, making visibility values cluster around some values (Fig. 3).

According to the experimental results and the cost of hyperspectral cameras, RGB images seem to be a better choice for environmental monitoring since RGB images can be easily and cheaply obtained through smartphones and webcams. At least for the weather stations, the RGB images can be used to complete the missing values caused by instrument malfunctions.

As the RMEP model only handles a single input image, it cannot grasp the temporal information between images. A combination of the RMEP model and some other sequence models will further improve the retrieval accuracy.

TABLE III
RETRIEVAL RESULTS ON THE TEST SET OF THE SSF DATASET UTILIZING DIFFERENT MODELS

Model	Visibility		RH		Temperature	
	R^2	MAE (m)	R^2	MAE (%)	R^2	MAE (°C)
AlexNet	0.62	1542.26	0.63	11.68	0.61	5.13
ResNet-50	0.61	1591.72	0.71	10.09	**0.83**	**3.37**
DenseNet-121	0.61	1706.85	0.69	10.70	0.79	3.81
RMEP	**0.63**	**1190.26**	**0.72**	**9.85**	0.82	3.44

TABLE IV
ABLATION STUDY USING DIFFERENT POOLING LAYERS AND NUMBERS OF RESNET BLOCKS

	Visibility		RH		Temperature	
	R^2	MAE (m)	R^2	MAE (%)	R^2	MAE (°C)
Max pooling	0.63	**1190.26**	0.72	**9.85**	0.82	**3.44**
Average pooling	**0.64**	1221.88	0.72	9.90	0.82	3.49
#ResNet blocks: 0	**0.65**	1277.91	0.70	10.25	0.80	3.74
#ResNet blocks: 3	0.63	1344.94	0.72	9.87	0.81	3.52
#ResNet blocks: 6	0.63	**1190.26**	0.72	**9.85**	0.82	**3.44**
#ResNet blocks: 9	0.63	1207.02	0.72	10.03	0.82	3.51

IV. CONCLUSION

An end-to-end CNN called RMEP is proposed for the retrieval of multiple atmospheric environmental parameters including atmospheric horizontal visibility, relative humidity, ambient temperature, $PM_{2.5}$, and PM_{10}. Experimental results demonstrate that spatial and spectral resolutions of images are not the key factors for the retrieval on the horizontal scale.

In the future, we will try to combine CNNs and sequence models to obtain the spatio-temporal patterns of atmospheric environmental parameters, in addition to studying the possibility of retrieving other atmospheric environmental parameters.

ACKNOWLEDGMENT

The work was supported by the MegaSense research programme of the University of Helsinki, the City of Helsinki Innovation Fund, Business Finland, Business Finland Project 6884/31/2018 MegaSense Smart City, the European Commission through the Urban Innovative Action Healthy Outdoor Premises for Everyone (no. UIA03-240) and by the Urban Innovative Action HOPE - Healthy Outdoor Premises for Everyone (no. 47402211). This research was also financially supported by the Ministry of Science and Technology of the People's Republic of China (no. 2019YFC0214701), the National Natural Science Foundation of China (no. 41877306) and Beijing University of Chemical Technology. The authors wish to acknowledge CSC–IT Center for Science, Finland, for computational resources.

REFERENCES

[1] J. S. Apte, K. P. Messier, S. Gani, M. Brauer, T. W. Kirchstetter, M. M. Lunden, J. D. Marshall, C. J. Portier, R. C. Vermeulen, and S. P. Hamburg, "High-resolution air pollution mapping with google street view cars: exploiting big data," *Environmental science & technology*, vol. 51, no. 12, pp. 6999–7008, 2017.

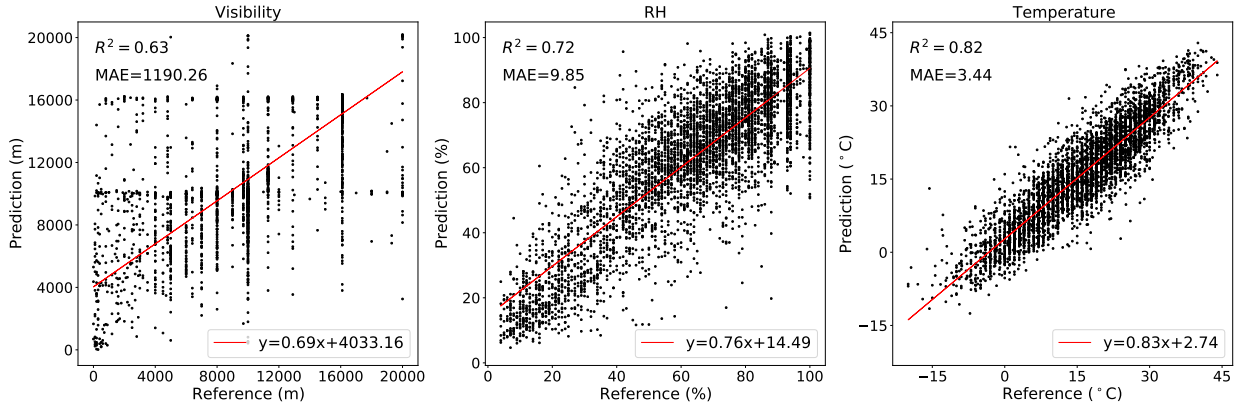

Fig. 3. Retrieval results on the test set of the SSF dataset.

[2] E. Lagerspetz, N. H. Motlagh, M. A. Zaidan, P. L. Fung, J. Mineraud, S. Varjonen, M. Siekkinen, P. Nurmi, Y. Matsumi, S. Tarkoma *et al.*, "Megasense: Feasibility of low-cost sensors for pollution hot-spot detection," in *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, vol. 1. IEEE, 2019, pp. 1083–1090.

[3] A. Nikandrova, K. Tabakova, A. Manninen, R. Väänänen, T. Petäjä, M. Kulmala, V.-M. Kerminen, and E. O'Connor, "Combining airborne in situ and ground-based lidar measurements for attribution of aerosol layers," *Atmospheric Chemistry and Physics*, vol. 18, no. 14, pp. 10 575–10 591, 2018.

[4] X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, "PM2.5 monitoring using images from smartphones in participatory sensing," in *2015 IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. IEEE, 2015, pp. 630–635.

[5] W.-T. Chu, K.-C. Ho, and A. Borji, "Visual weather temperature prediction," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018, pp. 234–241.

[6] Y. Yang, Z. Hu, K. Bian, and L. Song, "ImgSensingNet: UAV vision guided aerial-ground air quality sensing system," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 1207–1215.

[7] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492–505, 2018.

[8] L. K. Choi, J. You, and A. C. Bovik, "Referenceless perceptual fog density prediction model," in *Human Vision and Electronic Imaging XIX*, vol. 9014. International Society for Optics and Photonics, 2014, p. 90140H.

[9] ——, "Referenceless prediction of perceptual fog density and perceptual image defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888–3901, 2015.

[10] D. Glasner, P. Fua, T. Zickler, and L. Zelnik-Manor, "Hot or not: Exploring correlations between appearance and temperature," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3997–4005.

[11] A. Volokitin, R. Timofte, and L. Van Gool, "Deep features or not: Temperature and time prediction in outdoor scenes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 63–71.

[12] A. Chakma, B. Vizena, T. Cao, J. Lin, and J. Zhang, "Image-based air quality analysis using deep convolutional neural network," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3949–3952.

[13] Q. Zhang, F. Fu, and R. Tian, "A deep learning and image-based model for air quality estimation," *Science of the Total Environment*, vol. 724, p. 138178, 2020.

[14] X. Su, X. Liu, N. H. Motlagh, J. Cao, P. Su, P. Pellikka, Y. Liu, T. Petäjä, M. Kulmala, P. Hui *et al.*, "Intelligent and scalable air quality monitoring with 5G edge," *IEEE Internet Computing*, 2021.

[15] Y. You, C. Lu, W. Wang, and C.-K. Tang, "Relative CNN-RNN: Learning relative atmospheric visibility from images," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 45–55, 2018.

[16] L. Liu, W. Liu, Y. Zheng, H. Ma, and C. Zhang, "Third-eye: A mobilephoneenabled crowdsensing system for air quality monitoring," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–26, 2018.

[17] Z. Pan, H. Yu, C. Miao, and C. Leung, "Crowdsensing air quality with cameraenabled mobile devices." in *AAAI*, 2017, pp. 4728–4733.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[19] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," *arXiv preprint arXiv:1607.08022*, 2016.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[21] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.

[22] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.

[23] Y.-K. Wang and C.-T. Fan, "Single image defogging by multiscale depth fusion," *IEEE Transactions on image processing*, vol. 23, no. 11, pp. 4826–4837, 2014.

[24] N. Hautiere, J.-P. Tarel, J. Lavenant, and D. Aubert, "Automatic fog detection and estimation of visibility distance through use of an onboard camera," *Machine Vision and Applications*, vol. 17, no. 1, pp. 8–20, 2006.

[25] X. Wang, R. Zhang, and W. Yu, "The effects of PM2.5 concentrations and relative humidity on atmospheric visibility in beijing," *Journal of Geophysical Research: Atmospheres*, vol. 124, no. 4, pp. 2235–2259, 2019.

[26] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang *et al.*, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sensing of Environment*, vol. 241, p. 111716, 2020.

[27] P. Su, J. Joutsensaari, L. Dada, M. A. Zaidan, T. Nieminen, X. Li, Y. Wu, S. Decesari, S. Tarkoma, T. Petäjä, M. Kulmala, and P. Pellikka, "New particle formation event detection with mask r-cnn," *Atmospheric Chemistry and Physics*, vol. 22, no. 2, pp. 1293–1309, 2022. [Online]. Available: https://acp.copernicus.org/articles/22/1293/2022/

[28] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[29] D. Engin, A. Genç, and H. Kemal Ekenel, "Cycle-dehaze: Enhanced cyclegan for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 825–833.

[30] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448.