Automated Air Pollution Detection and Estimation of AQI Using ResNET

R.Udaya Shanmuga[1], Dr.G.TamilPavai[2]
[1]*Ph.D Scholar, CSE, Government college of Engineering,Tirunelveli-7*
[2]*HOD, CSE Department, Government college of Engineering,Tirunelveli-7*

***Abstract* - Air Pollution is turning the whole world prematurely grey. The health effects of air pollution imperil human lives especially for at-risk population and those with respiratory illness and the fact is well-documented. Many researches proposed to estimate particulate matter values from smartphone images, given that deploying highly accurate air pollution monitors throughout a city is a highly expensive. Departing from previous machine learning studies which primarily focus on pollutant estimation based on single day-time images, our proposed deep learning model integrates Residual Network (ResNet) with Long Short-Term Memory (LSTM), extracting spatial-temporal features of sequential images taken from smartphones instead for estimating PM2.5 and PM10 values of a particular location at a particular time. Our methodology is as follows: First, images are obtained constantly with regular time intervals. Second, verified experimentally that any PM2.5 and PM10 values obtained remain constant within a radius of 500 meters. Third, the proposed ResNet-LSTM was constructed and extended by incorporating meteorological information and one short path. In future, our deep-learning image-based air pollution estimation study will incorporate sequential images obtained from 24-hr operating traffic surveillance cameras distributed across all parts to provide full- day and more fine-grained image-based air pollution estimation for the city.**

I.INTRODUCTION

PM2.5 and PM10 have presented great public health challenges given their devastating health impacts, especially for those who are constantly living under high of air pollution, such as China and India. Citizens in these countries are consistently exposed to high levels of air pollution due to rapid industrialization and urbanization. In response to this air pollution related health challenge, governments all over the world have set up regular stationary air quality monitoring systems. Stringent air quality regulatory standards are put forward and air pollution reports are provided to inform the public the level of air pollution on a regular basis. For example, Pollutant Standards Index (PSI) has been introduced in Singapore and Daily Air Quality Index DAQI) has been deployed in the United Kingdom (UK). Given its harmful health consequences, it is important to make PM2.5 and PM10 readings throughout the city publicly available so any citizens can plan their outdoor activities accordingly.

Due to the high cost of building and operating government-run AQMS, only a limited number of AQMSs can be provided throughout a city to provide regular air pollution readings. In HK, only 18 AQMSs across an area of 1100 sq. km are available [1]. People residing in areas without an AQMS can only rely on measurements obtained from nearest AQMSs, which may not necessarily reflect the actual pollution readings of their own locations. To overcome such challenge, previous estimation studies designed and used feature-based machine learning models [7]–[11], such as Support Vector Regression (SVR), to estimate PM2.5 values, capitalizing on features manually selected from images (such as contrast and saturation). However, as these feature-based models are highly dependent on how features are constructed, their performance can be easily distorted by any change in environmental conditions. For example, night-time images normally have a lower intensity when compared to day-time images. Since image features are manually selected for air pollution estimation, conventional machine learning methods can be less robust when compared to deep learning-based methods, which can extract image features automatically In recent years, some studies [12]–[15] have used deep learning models to improve image estimation accuracy and performance. Although deep learning models have been able to extract image

features automatically, they have relied only on day-time images as inputs. Night-time image estimation remains a challenge, given that the low image intensity during night-time has often resulted in poor estimation accuracy. Furthermore, these studies have only considered spatial features extracted from single images via CNN models, while temporal correlations of sequential images are ignored. This is because CNN models can only learn features from single images, but not how images are changed from one temporal instance to the next, and recurrent patterns such as day-time vs. night-time images. To improve the quality of image-based PM estimation using night-time images, an end to-end ResNet-LSTM model is proposed in this study. Our proposed model can estimate both PM2.5 and PM10 values directly from pictures taken by smartphones. It achieves lower estimation errors and possess a higher pollutant estimation capability, after incorporating both day-time and night-time images, when compared to other baselines such as CNN-based models. Main novelties of this image-based pollutant estimation study cover the following:

- A temporally fine-grained image set is constructed with corresponding PM2.5 and PM10 values labelled.
- 3024 images have been taken consecutively in a time- sequential covering the real day-time images.
- An end-to-end ResNet model using sequential images as inputs is constructed and achieves the best PM2.5 and PM10 estimation, when compared to state- ofthe-art baselines.
- Apart from taking single images as inputs, ResNet- also incorporates sequential images taken once every minute for estimating PM2.5 and PM10 values of a specific location to enhance robustness.
- First deep learning model used to estimate PM2.5 and PM10 values based on night-time images.
- A novel Met-ResNet model is developed based on the newly developed ResNet model, taking into account six meteorological features, in addition to images taken from smartphones, cctv as inputs, which gives even better estimation performance when compared with the ResNet model.
- After incorporating deep supervising techniques, ResNet and Met-ResNet are constructed to improve estimation performance by comparing results of each with that of ResNet and Met-ResNet,

The rest of this paper is organized as follows. Related studies are reviewed in Section II. The Priciple is described in III. The dataset and the methodology are described in Section IV and V. Section VI presents both our experimental results and analyses, and Section VII concludes and highlights directions for future study.

II. RELATED WORK

This section reviews deep learning techniques and methods used for PM2.5 estimation, based on images.

A.DEEP CONVOLUTION NEURAL NETWORK:

Krizhevsky et al. [16] first constructed a deep convolutional neural network (CNN), AlexNet, to classify images into different categories in 2012. Its accuracy is higher when compared to manually selected feature methods. Simonyan [17] further increased the depth of CNN by adding very small convolution kernels. This further improves the image classification performance based on deep learning. The deep residual network, ResNet [18], also a CNN, was developed to tackle the issue of degradation that occurred when the number of neural network layers was increased in 2016. CNN is good at extracting key information from images. The technique has been widely implemented in the field of computer vision, such as facial recognition [19], image classification [16]–[18] and visual tracking [20]. Zheng et al. [21] and Hong et al. [22] used CNN models to analyze satellite images and estimate ground-level PM2.5 values. Apart from CNN, Hochreiter and Schmidhuber [23] proposed an LSTM model to extract features from sequential data for neural machine translation [24], [25]. Using these techniques, efforts have been made on combining CNN with LSTM for extracting spatial-temporal features. Chen et al. [26] used CNN-LSTM models to forecast typhoon formation and hourly air pollution across the city [27], [28]. These models have shown that pollution estimation performance can be improved by combining temporal features with spatial features.

B. PM ESTIMATION USING MACHINE LEARNING:

PM2.5 can affect the light scattering coefficient [29] when a picture is taken, as it obscures the scene and

blurs the sky, which eventually degrades the visibility [30]. Estimating pollution level based on smartphone-taken images is handy, as this allows one to easily capture any change in air pollution level. Conventional machine learning approach maps ambient light scattering coefficients with PM2.5 values. The haze-image model [31] was widely utilized to estimate scattering coefficients from single images. This model learned the formation of observed images (haze-images) from pure scenes that carry light-scattering effects. To estimate PM2.5 values, some studies combined the dark channel model [32] with the haze-image model to calculate light coefficients directly from single images [33], and Yang and Chen [34] made good use of the relative humidity to improve pollution estimation. As compared to [33], [34], Liu et al. [7] and Zhang et al. [8] extracted image features such as image entropy, contrast, and saturation for further improvement. Capitalizing on the haze-image model and extracted image features, Liu et al. [7] deployed Support Vector Regression (SVR) to estimate PM2.5 concentrations. Zhang et al. [8] made good use of multi-kernel learning to estimate air quality. Liu et al. [9] adopted similar features and used a linear least square regression to estimate PM2.5 values via smartphone-taken images. Instead of using basic image features, Gu et al. [10] constructed a picture-based predictor. The entropy information from the image saturation map was extracted and non-linear mapping was used to estimate PM2.5 values based on the overall likelihood of naturalness. Yue et al. [11] combined the color information loss with the structural information loss and applied a five-parameter logistic function to estimate PM2.5 values, which achieved a high estimation performance. After all, the performance of pollution estimation based on the conventional machine learning approaches can be easily distorted due to changes in meteorological conditions and discrepancies of light intensity between the day and the night

C.DEEP LEARNING FOR PM2.5 AND PM10 ESTIMATION:

Capitalizing on strengths of the previous deep learning techniques, Li et al. [35] combined the dark channel and the haze-image models to estimate the scattering ability of a medium, also referred to as the transmission layer in the haze-image model. Following this, the depth of an image was further estimated by Deep Convolutional Neural Fields (DCNF) [36], and a non-linear mapping was designed to estimate PM2.5 values based on experimental results. Since CNN can extract important spatial features from images, efforts have been made to construct CNN models to analyze images directly, without selecting image features manually. CNN models were used to analyze images and estimated PM2.5 [12]–[14] and PM10 levels [12], [14]. To further improve classification accuracy, Ma et al. [37] and Wang et al. [38] developed two parallel CNN models to analyze single images. Ma et al. [37] used two parallel CNNs to analyze both the original image and the transmission layer, while outputs from these two CNNs were used to estimate PM2.5. Images collected by Wang et al. [38] covered both skies and buildings, using the same weights, similar to the method outlined in [13]. Wang et al. [38] split images into two parts, the sky and buildings, and constructed a Double-Channel CNN to estimate air quality. Instead of providing a rough estimation, Liu et al. [15] used Long Short-Term Memory (LSTM) [23] to analyze the meteorological data, and used a CNN model to process images. Results from these two were combined to estimate PM2.5 concentrations. However, the models did not take sequential images into account. The pollutant estimation models above usually took PM2.5 values from government AQMSs close to the place where images were taken, rather than the exact measurements of the place where images were taken, as ground truths [39]. Further, these models had yet considered how environmental variations will affect image estimation performance. Image-based estimation performance could be seriously degraded given that images taken and PM2.5 measurements were not co-located, and that PM2.5 concentrations could change when environmental conditions, such as the street canyon effect, urban morphology and traffic conditions, change overtime. To tackle the current research gaps, we used high accuracy calibrated portable pollutant sensors to provide accurate PM2.5 and PM10 measurements of the exact locations where pollutant pictures have been taken. Furthermore, we incorporated six most relevant features reflecting meteorological conditions of locations that these pictures were taken into our deep learning model. Furthermore, previous studies extracted image features, such as entropy, saturation, or high-level

features layers, from the depth and the transmission map. However, these features could be easily distorted by any change in environmental conditions and light intensity, especially features that are extracted from low-intensity images, such as night-time images.

In our study, we combined ResNet and LSTM to extract spatial-temporal domain features. In contrast to [15], in our proposed model, an LSTM model is added to the ResNet to extract the temporal features from images. As single night-time images are difficult to be analyzed directly, earlier efforts were only based on single day-time images as inputs. Instead of simply processing single images, our model considers the sequential day-time and night-time images in estimating PM2.5 and PM10 values of a specific location.

III. PRINCIPLE

PM in air affects an optical image in different ways, but they are all originated from the interactions of light with the airborne particles, mainly via light scattering, including Rayleigh scattering and Mie scattering. Light scattering causes an attenuation of light transmission in air,which can be expressed by the Beer-Lambert law,

$t = ��^{-\beta d}$ (1)

where β is the medium extinction coefficient, which depends on particle size and concentration, and d is the distance of light propagation. This equation indicates that if the extinction coefficients at different wavelengths are determined, then PM concentration can be estimated. The extinction coefficient may be determined from an observed image according to

$I(x,y)=t(x,y)+J(x,y)+(1-t(x,y))A$ (2)

where I is the observed hazy image, t is the transmission from the scene to the camera, J is the scene radiance, A is the airlight color vector (see explanation below). As shown in Fig 1, the first term of Eq 2 is the direct transmission of the scene radiance into the camera, which is lightreflected by the object surfaces in the scene and attenuated by air before entering the camera. The second term (1-t(x,y))A is called airlight, which is the ambient light scattered by air mole- cules and PM into the camera /Wang et al. [9] applied the above formula to estimate light attenuation. In the present work, the relationship between transmission value and PM density was evaluated by analyzing ROIs at difference distances. Eq 2 assumes constant atmospheric and lighting conditions, which, in practice, may both change with the weather and position of the sun that vary with the time of the day and season. Additionally, both J and A depend not only on the weather and position of the sun, but also on PM distribution and con- centration. The present work considers these varying factors as additional features to improve the accuracy of PM estimation based on images. Fig 1.

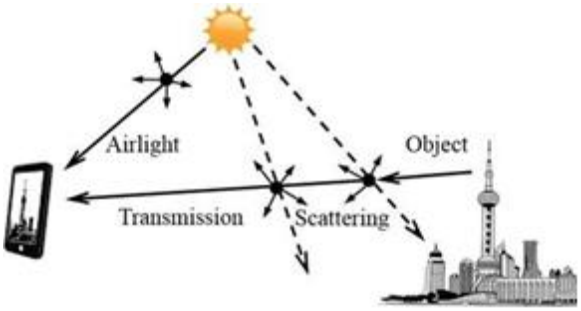

Fig 1. The radiance reaching the smartphone camera is the summation of the transmitted light fromthe object and airlight from the sun after scattering by air, water and PM in atmosphere.

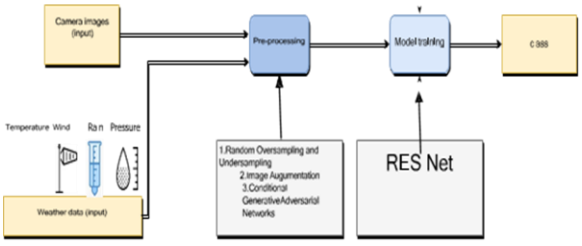

Fig: 2 PM estimation methodology using ResNet

The above discussion did not consider color information explicitly, which can also serve as important features for PM estimation based on light scattering consideration. Rayleigh scattering dominates when the particles (including air molecules) are much smaller than the wavelength of light. It is strongly wavelength dependent, and varies with wavelength (λ) according to λ-4, which is responsible for the blue color of the sky. In contrast, Mie scattering occurs when the size of the particles is comparable to the wavelengths of light, which tends to produce a white glare around the sun when particles are present in air. The combination of Rayleigh and Mie scattering affect the brightness and color saturation of an outdoor image. Conversely, the color and brightness information contains particle concentration and size information, and can be used as distinct features to estimate PM. The present work includes color

information as important image features for PM estimation, in addition to light attenuation

IV. METHODOLOGY

Our overarching methodology for estimating PM2.5 and PM10 values via sequential smartphone taken images consists of four stages (see Figure 2). First the collected data were pre-processed. Second image features were extracted from the images, which were used, together with other relevant data, such as the position of the sun, date, time, geographic information and weather conditions, to predict PM2.5 index. Third, our proposed ResNet-LSTM was further refined by incorporating meteorological features and one short path to exploit incorporating meteorological features and one short path to exploit the fullest potential of ResNet-LSTM for PM2.5 and PM10 pollutant estimation

A. Data Preprocessing:

Generally, high air pollution takes place in specific periods of the year, causing a high-class imbalance between the numbers of data instances available for different levels of air pollutions. To Predicted evaluate the capability and accuracy of PM estimation based on image analysis, it is critical to build a database. In the present work, we collected images, as well as the date and time of each image, PM2.5 index, weather data and geographic location from fixed scenes in cities from Archive of Many Outdoor Scenes dataset, captured every hour from 8:00 a.m. to 16:00 p.m. The weather dataof the three cities were obtained from Weather Underground (http://www.wunderground. com/) and Weather Spark (https://weatherspark.com/).

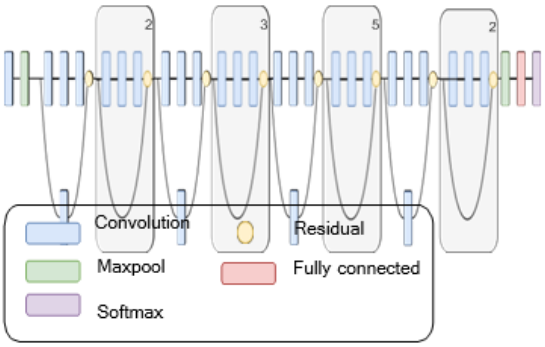

Fig:3 RESNet Architecture

Precise geographical locations,. Consequently, training our model using this type of dataset would ultimately lead the model biased to the non-polluted images. It would lead to over-fitting on the validation and testing set. Imbalanced datasets cause very impactful bottlenecks and need to be addressed. For this reason, we exploit random oversampling and undersampling as well as image augmentation, as explained in the following subsections.

B.Random Oversampling and Under sampling: This method used two different strategies for tackling the class imbalance issue. The first technique is random oversampling and under sampling, as explained. This technique allows obtaining a balanced dataset by randomly taking multiple copies of minority class images or randomly by skipping a portion of the majority class images. Tse techniques did not prove useful in our problem and did not resolve the imbalance issue.

C.Augmentation: Standard image processing augmentation techniques were applied recurring to the Keras library. Generating new and realistic images with operations such as image flipping and rotation, as well as zooming by small factors, has a big potential in improving the performances of the network. These techniques were applied after both approaches for class balancing described in the following subsections.

D.Conditional Generative Adversarial Networks (CGAN): Conditional Generative Adversarial Networks (CGAN) allows us to augment the dataset using generative models. Here trained these models to learn how to generate new images based on previously seen images. Specifically, we trained a CGAN model using the minority class's images in the training set and applied it to generate new images for the same class, mitigating imbalance issue. The idea of the CGAN is the first class, and the remaining are of the second class. The CGAN is not used in the testing dataset because we want to test our models on the camera's raw images. After the CGAN is applied and the dataset has been balanced; simple augmentation methods such as flipping images, rotating, and zooming by small factors are applied to the dataset. This additionally improves the performances of our models and provides us with a broader dataset to augment the dataset and provide us with a training dataset where 50% of the images are from.

V.RESNET ARCHITECTURE

The main benefit of a very deep network is that it can represent very complex functions. It can also learn features at many different abstraction levels, from edges (at the shallower layers, closer to the input) to very complex features (at the deeper layers, closer to the output). However, very deep neural networks are difficult to train because of the problem of vanishing gradient. The vanishing gradient problem occurs when the gradient is back-propagated to earlier layers, and the repeated multiplication operations make the gradient infinitely small. As a result, as network becomes deeper, its performance saturates or even degrades rapidly. This problem inspired the residual network (ResNet), which allows robust deep convolutional neural networks to be built. The main idea of the ResNet model is the residual blocks, whose architecture is shown in Figure 4. The residual block provides two paths for the input: the main path and the shortcut (or more commonly known as skip-path). The main path provides the normal flow as with any convolutional neural network. Still, the shortcut skips N convolution layers (in our approach N = 2) and provides its input to the following convolution layer. The authors in argue that stacking layers should not degrade the network performance because one could simply stack identity mappings (layers that learn the identity mapping which ultimately does not make any change) upon the current network, and the resulting architecture would perform equally. This indicates that the deeper model should not produce a training error higher than its shallower counterparts. They hypothesize that letting the stacked layers fit a residual mapping is easier than letting them directly fit the desired underlying mapping. The residual block explicitly allows the model to perform this task .The basic residual block is also known as the identity block. This block corresponds to the case where the input activation has the same dimension as the output activation. Apart from this block, there is another block called the convolutional block. The architecture of this block is shown in Figure 5. In this block, the input and output dimensions do not match. The difference between this block and the identity block is that there is a convolutional layer in the shortcut path to match the dimensions

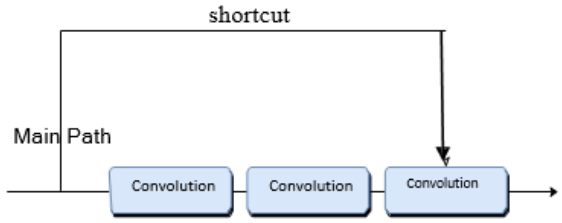

Fig:4 Basic Residual Block

The complete architecture of the ResNet model (Figure 3) is that the model is built using multiple convolutional and identity blocks. It also contains convolution and max pooling layers at the beginning of the model, as well as flatten and dense layers at the end to make the final prediction. By using this model, better results can be obtained which can be explained by the fact that much deeper networks can be used that can learn complex features and are not affected by the problem of vanishing gradient.

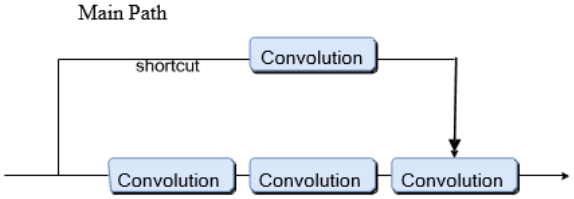

Fig: 5 Convolution Residual block

Conventional pollution estimation method adopts a haze- image model, which is integrated with the dark channel prior techniques. The haze-image model developed by shows how the haze image can beoverlaid by pure sceneries and scattering effects. This model has image dehazing and light scattering coefficient estimation, as shown in (3) below: $I(x)=Ot(x)+A(1-t(x)$ (3) where I(x) represents the observed scenery captured by the camera, O represents the pure scenery without attenuation, A represents the atmospheric light, and t(x) is the scattering ability of the medium over the light path length. The transmission t(x) can be expressed as $e-\beta x$, where β is the scattering coefficient of the atmosphere and next, the ambient scattering coefficient β is calculated foreach single pixel, and the average β is calculated for the entire image. Finally, based on the estimated ambient scattering coefficients, linear regression models are used to estimate the PM2.5 and PM10 values.

VI. ESTIMATION OF AQI

For conducting our experiments and evaluation, we used Tensorflow version 2.3.0 and Kerasversion 2.4.3. The complete dataset consists of 178,992 images, where 80% of the images are taken for training the models, 15% of the images are for testing, and 5% for validating. We decided to use a smaller percentage for validation and testing datasets because we have a bigger dataset, and using this approach, we can leverage the large amount of data for training purposes. The validation dataset was used for hyper-parameter tuning, where a grid search approach was used. The hyper-parameters tuned are: learning rate, batch size, and optimizer used. We have obtained the optimal parameters through the grid search, i.e., a learning rate of 0.001, batch size of 64 images, and Adam optimizer. Our proposed models have incorporated pollution data and weather information for the last two years .The air pollution data is collected from the API endpoints of pulse.eco (https://pulse.eco). The API provides information about different sensors and different pollutants at different timesteps. The timesteps are irregular, so the data is aggregated hourly for each sensor. Every row consists of weather information, timestamp (date and time), the pollutant type, and the amount measured. Our models have used only the PM2.5 pollution to labeling the images during the training process. Therefore, they are not inputs in the machine learning models. This method also uses images taken from a stationary camera .The camera takes periodical pictures of the center of the city, and at the same time, air pollution sensors measure the exact air quality. Based on the air quality measurements in terms of PM2.5 concentration, the images were labeled with six classes depending on the Air Quality Index (AQI) of the European Union, as shown in Table 1.The implications of the six AQI indexes are the following. AQI-1 means that the air quality is satisfactory, and air pollution poses little or no risk. AQI-2 means that the air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution. AQI-3 means that members of sensitive groups may experience health effects. The general public is less likely to be affected. With AQI-4, some members of the general public may experience health effects, and members of sensitive groups may experience more serious health effects. AQI-5 requires a health alert because the risk of health effects is increased for everyone. AQI-6 entails issuing a health warning of emergency conditions because everyone is more likely to be affected. Ideally, a system would be able to estimate air pollution precisely. However, the fact that the proposed models are attempting to do this based on the camera images and not actual air pollution sensors makes it unreasonable to define the task as a regression problem. Therefore, our initial approach was to use six classes, one for each AQI category. We additionally redefined the problem as a binary classification problem, collapsing AQI-1 and AQI-2 into the "not polluted" class, and the otherAQI indexes into the "polluted" class. In some models, we have also incorporated the weather information to distinguish between weather conditions and pollution. The weather information was collected from the API endpoints of World Weather Online (https:/ /www.worldweatheronline.com). The data consists of temperature, wind speed, wind direction, weather description, precipitation, humidity, visibility, pressure, cloud coverage, heat index, and the UV index. This is used a simple strategy that considers the last measured value to handle the missing pollution measurements, although more sophisticated approaches based on generative models can be incorporated Even though we have merged the six classes into two general classes, the dataset is still highly imbalanced, and our models could become very biased to non-polluted images. For that purpose, we applied different techniques for balancing the dataset and evaluated their impact on the classification performance.

AQI Category	PM2.5Range	6-Class Labels	Binary Labels
Good	0–50	AQI-1	Not polluted
Moderate	51–100	AQI-2	Not polluted
Unhealthy for Sensitive Groups	101–150	AQI-3	Polluted
Unhealthy	151–200	AQI-4	Polluted
Very Unhealthy	201–300	AQI-5	Polluted
Hazardous	301and above	AQI-6	Polluted

Table 1. Air Quality Index (AQI) categories based on

predictions are not useful, which was one of the main reasons to evaluate the binary classification approach that collapsed multiple categories into only two .The training and testing accuracy of the different architectures, depending on the number of epochs and whether class balancing was performed or not is predicted accurately. These results confirm that the

models are stable, and after about 40 epochs, the performance does not vary. Likewise, it is evident that both that training and testing accuracy benefited significantly from the proposed CGAN data augmentation. Note that the test set remains unbalanced in all experiments because the balancing is performed only on the training set. The reason for that is that in a production setting, the camera images would not be going that class over the total number of images

VII.LIMITATIONS AND FUTURE WORK

This study aims to estimate PM2.5 and PM10 concentrations of a specific location in HK using sequential images taken consecutively once every minute, addressing also the night-time image estimation challenge. Our results show that the proposed ResNet and the Met-ResNet model can achieve a better PM2.5 and PM10 estimation when compared to the baselines, especially after incorporating the short path. However, a few limitations remain.

First, although deep-learning models can learn the most important features that are proxies to PM2.5 and PM_{10} pollutant concentrations, the estimation performance can still be affected by other pollutants,

ranges of PM2.5 values and mapping to labels in Table 2 shows the distribution of the dataset in when using six classes. The distribution of the dataset after collapsing the six classes into two for illustrative purposes.

Dataset	AQI-1	AQI-2	AQI-3	AQI-4	AQI-5	AQI-6
Train	80,331	21,623	13,954	11,087	9862	6337
%	56.1%	15.1%	9.7%	7.7%	6.9%	4.4%
Test	13,342	4219	3022	987	1564	1135
%	52.8%	16.7%	12.0%	7.9%	6.2%	4.5%

This shows the different architectures' training and testing accuracy, depending on the number of epochs. It is clearly visible that the models are stable, and after a small number of epochs, the performance does not vary significantly. However, another clear result is that the accuracy is about 56%, the same as the majority class ratio. This means that the models learned to classify all images as "AQI-1," i.e., the "good" AQI category. As such, the as high O_3 and NO_2 values may also degrade the visibility. It is noted that image visibility can be affected by other pollutants, such as O_3, NO_2, and humidity. In the future, we can develop a model based on the same methodology to estimate values of other pollutants, such as O_3 and NO_2, based on sequential images taken.

VIII. CONCLUSION:

An end-to-end ResNet-LSTM model has been proposed to estimate PM2.5 and PM10 values from smartphone-taken images directly. Reliable estimation can be obtained for both day-time and night-time images. Our study consists of four stages. First, we have calibrated two low-cost portable sensors to provide reliable high accuracy pollutant measurements. Second, we have conducted an experiment to show that PM measurements within a distance of up to 500 meters are nearly constant. Third, based on our features and the empirical experiment, a comprehensive dataset containing 3024 images have been constructed. It has covered day-time images of the same building (up to 500 meters away), with all images taken by a camera. Finally, our proposed ResNet model works well and efficiently predicts particulate matters.

REFERENCES

[1] Environmental Protection Department. Accessed: Oct. 29, 2020. [Online]. Available: https://www. aqhi.gov.hk/en/what-is- aqhi/about-aqhi.html

[2] H. Akimoto, Atmospheric Reaction Chemistry (Springer Atmospheric Sci- ences). Tokyo, Japan: Springer, 2016.

[3] M. Kampa and E. Castanas, ''Human health effects of air pollution,'' Environ. Pollut., vol. 151, no. 2, pp. 362–367, Jan. 2008.

[4] X. Meng, Y. Zhang, K.-Q. Yang, Y.-K. Yang, and X.-L. Zhou, ''Potential harmful effects of PM2.5 on occurrence and progression of acute coronary syndrome: Epidemiology, mechanisms, and prevention measures,'' Int. J. Environ. Res. Public Health, vol. 13, no. 8, p. 748, Jul. 2016.

[5] J. C. Chow, J. G. Watson, J. L. Mauderly, D. L. Costa, R. E. Wyzga, S. Vedal, G. M. Hidy, S. L. Altshuler, D. Marrack, J. M. Heuss, G. T. Wolff, C. A. Pope, III, and D. W. Dockery, ''Health effects of fine particulate air pollution: Lines that connect,'' J. Air Waste Manage. Assoc., vol. 56, no. 10, pp. 1368–1380, Oct. 2006.

[6] Y. Xing, Y. Xu, M. Shi, and Y. Lian, ‘‘The impact of PM2. 5 on the human respiratory system,’’ J. Thoracic Disease, vol. 8, no. 1,p. E69, 2016.

[7] C. Liu, F. Tsow, Y. Zou, and N. Tao, ‘‘Particle pollution estimation based on image analysis,’’ PLoS ONE, vol. 11, no. 2, Feb. 2016, Art. no. e0145955.

[8] Z. Zhang, H. Ma, H. Fu, L. Liu, and C. Zhang, ‘‘Outdoor air quality level inference via surveillance cameras,’’ Mobile Inf. Syst., vol. 2016, pp. 1–10,Jan. 2016.

[9] X. Liu, Z. Song, E. Ngai, J. Ma, and W. Wang, ‘‘PM2: 5 monitoringusing images from smartphone

[10] Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2015, pp. 630–635.

[11] K. Gu, J. Qiao, and X. Li, ‘‘Highly efficient picture-based prediction of PM2.5 concentration,’’ IEEE Trans. Ind. Electron., vol. 66, no. 4,pp. 3176–3184, Apr. 2019.

[12] G. Yue, K. Gu, and J. Qiao, ‘‘Effective and efficient photo-based PM2.5 concentration estimation,’’ IEEE Trans. Instrum. Meas., vol. 68, no.10, pp. 3962–3971, Oct. 2019.

[13] C. Zhang, J. Yan, C. Li, H. Wu, and R. Bie,‘‘End-to-end learning for image-based air quality level estimation,’’ Mach. Vis. Appl., vol. 29, no. 4,pp. 601– 615, May 2018.

[14] A. Chakma, B. Vizena, T. Cao, J. Lin, and J. Zhang, ‘‘Image-based air quality analysis using deep convolutional neural network,’’ in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 3949–3952.

[15] C. Zhang, J. Yan, C. Li, X. Rui, L. Liu, and R. Bie, ‘‘On estimating air pollution from photos using convolutional neural network,’’ in Proc. ACM Multimedia Conf., 2016, pp. 297–301

[16] Liu, W. Liu, Y. Zheng, H. Ma, and C. Zhang, ‘‘Third-eye: A mobilephone-enabled crowdsensing system for air quality monitoring,’’ Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol., vol. 2, no. 1,pp. 1–26, 2018.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017.

[18] K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ 2014, arXiv:1409.1556. [Online]. Avail- able: http://arxiv.org/abs/1409.1556

[19] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[20] O. M. Parkhi, A. Vedaldi, and A. Zisserman, ‘‘Deep face recognition,’’ in Proc. Brit. Mach. Vis. Conf., BMVA Press, 2015, p. 41, doi: 10.5244/C.29.41.

[21] H. Nam and B. Han, ‘‘Learning multi-domain convolutional neural net- works for visual tracking,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4293–4302.

[22] T. Zheng, M. H. Bergin, S. Hu, J. Miller, and D. E. Carlson, ‘‘Estimating ground-level PM2.5 using micro-satellite images by a convolutional neural network and random forest approach,’’ Atmos. Environ., vol. 230, Jun. 2020, Art. no. 117451.

[23] K. Y. Hong, P. O. Pinheiro, and S. Weichenthal, ‘‘Predicting global variations in outdoor PM2.5 concentrations using satellite images and deep convolutional neural networks,’’ 2019, arXiv:1906.03975. [Online]. Available: http://arxiv.org/abs/1906.03975

[24] S. Hochreiter and J. Schmidhuber, ‘‘Long short-term memory,’’ Neural Comput., vol. 9, no. 8, pp.1735–1780, 1997.

[25] I. Sutskever, O. Vinyals, and Q. V. Le, ‘‘Sequence to sequence learning with neural networks,’’ in Proc. Adv. Neural Inf. Process. Syst., 2014,pp. 3104–3112.

[26] R. Chen, X. Wang, W. Zhang, X. Zhu, A. Li, and C. Yang, ‘‘A hybrid CNN-LSTM model for typhoon formation forecasting,’’ GeoInformatica, vol. 23, no. 3, pp. 375–396, Jul. 2019.

[27] C.-J. Huang and P.-H. Kuo, ‘‘A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities,’’ Sensors, vol. 18, no. 7, p. 2220,Jul. 2018.

[28] Q. Zhang, J. C. Lam, V. O. Li, and Y. Han, ‘‘Deep-AIR: A hybrid CNN-LSTM framework forFine-grained air pollution forecast,’’ 2020, arXiv:2001.11957. [Online]

[29] H. Ozkaynak, A. D. Schatz, G. D. Thurston, R. G. Isaacs, and R. B. Husar, ‘‘Relationships between

aerosol extinction coefficients derived from air-port".

[30] "Visual range observations and alternative measures of airborne particle mass,'' J. Air Pollut. Control Assoc., vol. 35, no. 11, pp. 1176–1185,Nov. 1985.

[31] N. P. Hyslop, ''Impaired visibility: The air pollution people see,'' Atmos. Environ., vol. 43, no. 1, pp. 182–195, Jan. 2009.

[32] S. K. Nayar and S. G. Narasimhan, ''Vision in bad weather,'' in Proc. 7th IEEE Int. Conf. Comput. Vis., vol. 2. 1999, pp. 820–827.

[33] K. He, J. Sun, and X. Tang, ''Single image haze removal using dark channel prior,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 12, pp. 2341–2353, Dec. 2011.

[34] H. Wang, X. Yuan, X. Wang, Y. Zhang, and Q. Dai, ''Real-time air quality estimation based on color image processing,'' in Proc. IEEE Vis. Commun. Image Process. Conf., Dec. 2014, pp. 326–329.

[35] B. Yang and Q. Chen, ''PM2.5 concentration estimation based on image quality assessment,'' in Proc. 4th IAPR Asian Conf. Pattern Recognit. (ACPR), Nov. 2017, pp. 676–681.

[36] Y. Li, J. Huang, and J. Luo, ''Using user generated online pho-tos to estimate and monitor air pollution in major cities,'' in Proc. 7th Int. Conf. Internet Multimedia Comput. Service (ICIMCS), 2015, pp. 1–5.

[37] F. Liu, C. Shen, G. Lin, and I. Reid, ''Learning depth from single monocular images using deep convolutional neural fields,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 10, pp. 2024–2039, Oct. 2016.

[38] J. Ma, K. Li, Y. Han, and J. Yang, ''Image-based air pollution estimation using hybrid convolutional neural network,'' in Proc. 24th Int. Conf. Pat- tern Recognit. (ICPR), Aug. 2018, pp. 471–476