

Research Article

Estimation of Particulate Levels Using Deep Dehazing Network and Temporal Prior

SeHee Jung ¹, SungMin Yang,² Eunseok Lee,¹ YongHak Lee,² Jisun Ko,¹ Sungjae Lee,¹ JunSang Cho,³ Jaehwa Lee,^{4,5} and SungHwan Kim ²

¹AI Analytics Team, DeepVisions, Seoul, Republic of Korea

²Department of Applied Statistics, Konkuk University, Seoul, Republic of Korea

³Industry University Cooperation Foundation Konkuk University, Seoul, Republic of Korea

⁴Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

⁵NASA Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence should be addressed to SungHwan Kim; shkim1213@konkuk.ac.kr

Received 7 May 2020; Revised 2 June 2020; Accepted 11 June 2020; Published 7 July 2020

Academic Editor: Bin Gao

Copyright © 2020 SeHee Jung et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Particulate matters (PM) have become one of the important pollutants that deteriorate public health. Since PM is ubiquitous in the atmosphere, it is closely related to life quality in many different ways. Thus, a system to accurately monitor PM in diverse environments is imperative. Previous studies using digital images have relied on individual atmospheric images, not benefiting from both spatial and temporal effects of image sequences. This weakness led to undermining predictive power. To address this drawback, we propose a predictive model using the deep dehazing cascaded CNN and temporal priors. The temporal prior accommodates instantaneous visual moves and estimates PM concentration from residuals between the original and dehazed images. The present method also provides, as by-product, high-quality dehazed image sequences superior to the nontemporal methods. The improvements are supported by various experiments under a range of simulation scenarios and assessments using standard metrics.

1. Introduction

Particulate matters (PM) are small particles suspended in the air, generally having an aerodynamic diameter smaller than or equal to $10\ \mu\text{m}$ (micrometers). PM originates from anthropogenic activities, (e.g., combustion of fossil fuel, dust) as well as natural sources (e.g., mineral dust, volcanic ash). PM measurements are commonly made for particles with an aerodynamic diameter smaller than or equal to $2.5\ \mu\text{m}$ ($\text{PM}_{2.5}$) and $10\ \mu\text{m}$ (PM_{10}). The size of PM is directly associated with health problems [1]. Inhaling the small particles is known to be hazardous as they can infiltrate in depth into the respiratory system [2]. In this regard, $\text{PM}_{2.5}$ has been widely used as a key indicator of the air quality index, and thus, we focus on $\text{PM}_{2.5}$ in this study (hereafter abbreviated as PM for brevity). Many experts point out that the recent increases of PM in many parts of the world are attributed to the rapid

growth in global energy consumption [3]. Over the years, efforts have been made to identify adverse effects of PM on public health and environment [4–6]. Surprisingly, the large-scale retrospective cohort study of lung cancer by the World Cancer Institute reported that PM is ascertained as primary carcinogens, as the risk of lung cancer increased by 22% for an increase of PM by $10\ \mu\text{g}/\text{m}^3$ [7]. The air pollution from PM has been one of the most controversial issues over East Asia. This has become no longer negligible, and media as well as researchers try to inform the public of its detrimental effects [8]. Particularly, it was reported that the annual mean PM concentration in South Korea is twice as high as in the Organisation for Economic Co-operation and Development (OECD; <https://www.oecd.org/>) countries [9]. Under this circumstance, it is imperative to build an accurate air monitoring system to facilitate alerting the public prevention. The Korean government drastically expanded the PM monitoring

network to help improve PM forecast service to a remarkable extent. However, there are still not enough PM monitoring stations to cover the whole country. Moreover, most of the stations are distributed in urban areas (e.g., Seoul, Busan), leaving many suburban and rural undermonitored.

Two different types of approaches can be used to estimate PM concentrations: sensor-based approaches and vision-based approaches.

1.1. Sensor-Based. Improvements in PM measurements using sensor-based approaches have been made to develop more precise sensor units [10]. There are two types of devices [11], i.e., microbalance PM monitoring stations (accurate but expensive [12]) and portable light-scattering-based PM monitors [13, 14]. For instance, the Korea Meteorological Administration (KMA) now operates 475 measuring stations and publicly reports PM concentration levels (i.e., limited to the station vicinity) every hour [15]. Most of the instruments operated by KMA are Tapered Element Oscillating Microbalance (TEOM), which is designed to directly weight PM on the filter [11]. Although highly accurate, TEOM is relatively expensive to install and maintain (approximately 200K USD per year) [16] and is bound to space limitations, thereby undermining practicality. The light-scattering method is relatively affordable. In [11, 17], collected PM via airflow structure are measured by densely deployed sensors. However, both methods relying on large-scale sensing nodes inevitably suffer from the expensive maintenance costs for high coverage and reliability. Recently, several newly developed devices for mobile facility (e.g., balloons and drones) are very interesting [10, 18, 19], but carrying sensors to acquire data is still highly energy-intensive and less practical to use.

1.2. Vision-Based. This approach is less explored compared to the sensor-based approach, and so there is a lot of room for improvement. To the best of our knowledge, all the vision-based studies exploit only individual images (e.g., [20–22]). In this case, it is fairly sensitive to motion blurring frequently caused by camera or subject movements. Specifically, Liu et al. [20] manually determined several regions of interest targeting distant objects to derive the transmission map. The explanatory power of the transmission map for PM estimation has proven its efficiency [20–23]. However, the need for selecting regions of interest is the pain point. Li et al. [21] leveraged heterogeneous data composed of GPS, camera lens, magnetic sensor, official station, and image data. Combining these multiple data, they generate high-dimensional features using kernel methods. Pan et al. [22] extracted haze effects using Adaptive Transmission Map [24] and pass derived features to the deep neural network designed on the basis of the well-known Boltzmann machine [25]. Since it is assumed that transmission values in a local patch (a.k.a. window) are the same constant, the dehazed images inevitably contain blocking artifacts [22, 26, 27]. The rest of this paper is organized as follows. In Sections 2 and 3, we introduce related works and the proposed method, respectively. Section 4 describes the numerical experiments.

We discuss the results and conclude in Section 5.

2. Related Work

According to the atmospheric scattering model (ASM), two factors are involved in the formation of a haze image: (1) direct attenuation and (2) airlight. When we take a photo, reflected radiance coming from objects is attenuated while reaching the camera. This is due to the effects of atmosphere absorption (a.k.a. direct attenuation), and the intensity of attenuation is proportional to camera distance (scene depth). In addition to this, there is another light, called airlight, resulting from scattering of neighboring light sources (e.g., sun) by haze [28]. Importantly, airlight is known to shift the color range of the object, and direct attenuation describes the scene radiance and its decay. Figure 1 illustrates how image degradation occurs under haze conditions. With a little of algebra, this process can be formulated as follows:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + \mathbf{A}(1 - t(\mathbf{x})), \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^2$ denotes the pixel, $\mathbf{I} \in \mathbb{R}^3$ is the observed haze image in RGB channels, $\mathbf{J} \in \mathbb{R}^3$ is the haze-free image, $t \in [0, 1]$ is the medium transmission describing the portion of the scene radiance that reaches the camera, and \mathbf{A} is the global atmospheric light. Note that \mathbf{A} is assumed to be homogeneous throughout the image and determined using empirical techniques. The first term $\mathbf{J}(\mathbf{x})t(\mathbf{x})$ on the right-hand side of equation (1) represents direct attenuation, and the second term $\mathbf{A}(1 - t(\mathbf{x}))$ corresponds to airlight. The farther the distance between a camera and objects is, the thicker the atmospheric layer exists between them.

$$t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}, \quad (2)$$

where β is the scattering coefficient of atmosphere, and $d(\mathbf{x})$ is the distance from the object to the camera. Equation (2) indicates that the scene radiance is attenuated exponentially with the scene depth d . Here, it is intuitively understandable that some haze effect $\mathbf{H}(\mathbf{x})$ deteriorates the haze-free image $\mathbf{J}(\mathbf{x})$ (i.e., $t(\mathbf{x}) = 0$), causing the haze image $\mathbf{I}(\mathbf{x})$. Starting from this simple idea, we can consider the following relation:

$$\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x}) + \mathbf{H}(\mathbf{x}). \quad (3)$$

Related to image dehazing (a.k.a. haze removal), we restore \mathbf{J} , \mathbf{A} , and t from the observed haze image \mathbf{I} .

3. Proposed Method

To measure PM via image sequences (i.e., video clip) is the ultimate goal of our study, and the deep neural network serves as a feature extractor. To achieve our goal, we first learn the deep dehazing network that extracts informative features based on strong correlation with PM concentration levels. In this section, we introduce two strategies working together for the network: (1) deep compression for energy efficiency in light of the model architecture and (2) temporal priors to capture haze-related features in image sequences. The specific design of our dehazing network is presented in Figure 2.

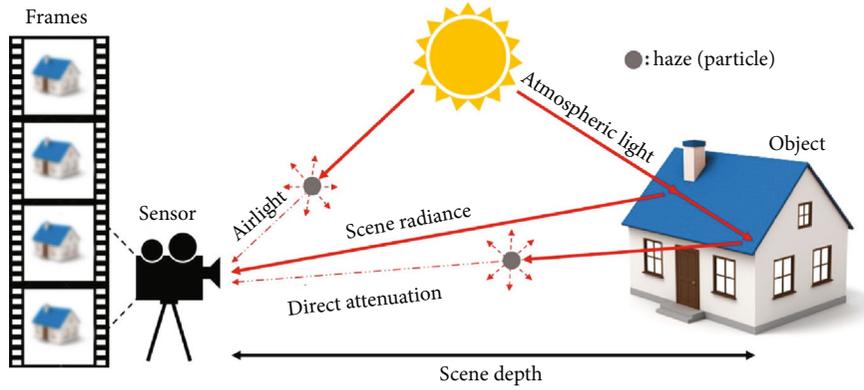


FIGURE 1: Hazy objects in multiple image sequences via the atmospheric scattering model.

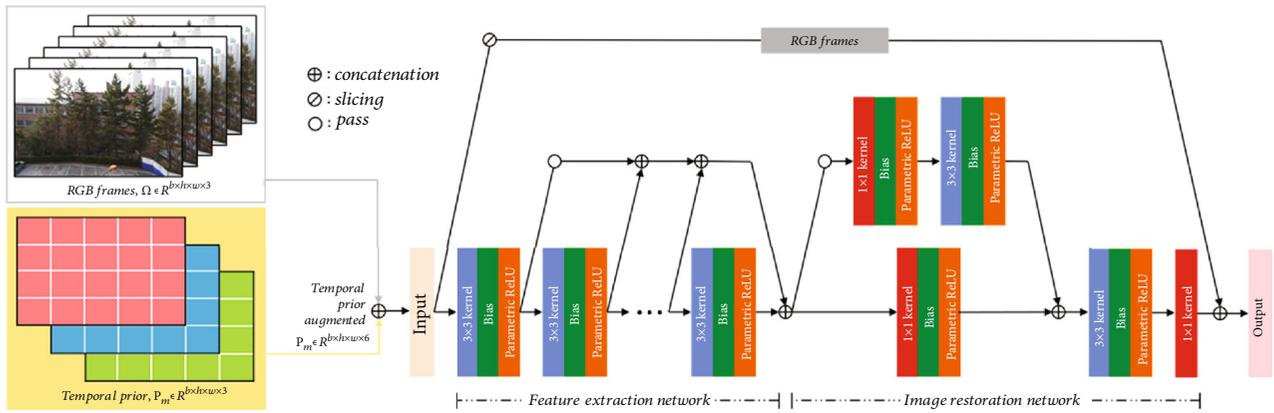


FIGURE 2: The illustration of dehazing network architecture. The dehazing network consists of two parts with specific roles: (1) feature extraction and (2) image restoration.

3.1. Network Architectures in Favor of Feature Extraction

3.1.1. Network Pruning. Network depth in CNN plays a decisive role in extracting various levels of features [29]. Inspired by this, we formulate the feature extraction network (FEN) at the front part of our dehazing network with deep cascaded convolutional layers of 436 kernels as in Figure 2. However, going deep convolutions involves too many parameters in the network and high computational cost. We thus seek to simplify the network without loss of accuracy. To address this hurdle, we compress the FEN using network pruning (NP). This has been widely used to reduce network complexity and prevent overfitting [30–32]. We also eliminate a waste of unnecessary parameters in our network without performance loss. Since the local feature is more important than the global feature in image restoration [33], we reduced the number of kernels in a row (see Table 1). This technique (a.k.a. pruned network) outperforms the original network having far more connections with plain and dense architecture. This in turn enabled the FEN to be less complex, but still deep enough to extract the necessary features for PM measurements. Based on the deep cascaded CNN architecture, the FEN grasps all the levels of features from local to global.

3.1.2. Parallelized 1×1 CNN Layer. In the second place, we adopt in the network the parallelized 1×1 CNN Layers

(P1CL, a.k.a. Network In Network [34]) to image restoration network (IRN) that appears in the second half of our dehazing network in Figure 2. The P1CL proves its applicability to enhance the representational power of neural networks [34]. GoogLeNet fully exploits this technique not only to make the network deep in the name of the inception module but also to reduce the dimensions inside these modules [35]. Inspired by GoogLeNet, we place the P1CL in front of the IRN to reduce the dimension of the feature maps accumulated through the FEN. This does not simply imply more of dimension reduction. In doing so, we compress the feature maps so as to narrow down the scope of involving features. Besides, the P1CL also includes the use of rectified linear activation [36]. In general, each P1CL consists of one or more 1×1 convolutions followed by a nonlinear activation function (e.g., ReLU), which adds more nonlinearity and thereby helps approximate the highly nonlinear function such as ASM. Taking all together, we form the IRN approximating the following equation derived from equation (3) with a little algebra:

$$\mathbf{H}(\mathbf{x}) = \mathbf{J}(\mathbf{x})(t(\mathbf{x}) - 1) + \mathbf{A}(1 - t(\mathbf{x})), \quad (4)$$

where $\mathbf{H}(\mathbf{x})$ means the haze effect and the two terms $\mathbf{J}(\mathbf{x})(t(\mathbf{x}) - 1)$ and $\mathbf{A}(1 - t(\mathbf{x}))$ correspond to the lost scene radiance (e.g., by scattering or absorbing) and the shift

TABLE 1: The architecture of the proposed dehazing network. In the table, b , h , and w represent batch size, height, and width, respectively. *concat* is short for concatenation.

Layer	Kernel	Input size	Output size
fen_1	$3 \times 3 \times 96$	$b \times h \times w \times 6$	$b \times h \times w \times 96$
fen_2	$3 \times 3 \times 83$	$b \times h \times w \times 96$	$b \times h \times w \times 83$
fen_3	$3 \times 3 \times 72$	$b \times h \times w \times 83$	$b \times h \times w \times 72$
fen_4	$3 \times 3 \times 61$	$b \times h \times w \times 72$	$b \times h \times w \times 61$
fen_5	$3 \times 3 \times 51$	$b \times h \times w \times 61$	$b \times h \times w \times 51$
fen_6	$3 \times 3 \times 41$	$b \times h \times w \times 51$	$b \times h \times w \times 41$
fen_7	$3 \times 3 \times 32$	$b \times h \times w \times 41$	$b \times h \times w \times 32$
<i>concat</i>	—	—	$b \times h \times w \times 436$
irn_1	$1 \times 1 \times 64$	$b \times h \times w \times 436$	$b \times h \times w \times 64$
irn_{21}	$1 \times 1 \times 32$	$b \times h \times w \times 436$	$b \times h \times w \times 32$
irn_{22}	$3 \times 3 \times 32$	$b \times h \times w \times 32$	$b \times h \times w \times 32$
<i>concat</i>	—	—	$b \times h \times w \times 96$
irn_3	$3 \times 3 \times 3$	$b \times h \times w \times 96$	$b \times h \times w \times 3$
irn_4	$1 \times 1 \times 3$	$b \times h \times w \times 3$	$b \times h \times w \times 3$

of the scene color, respectively. The main reason for parallel processing in IRN is to divide the computational tasks and thereby reduce the burden of the network. The upper and lower unit estimates the lost scene radiance and the shift of the scene, respectively. In the blessing of a parallel architecture, the IRN can expedite the work in a simultaneous processing fashion, and it also helps prevent overfitting problem with reduced network complexity. The residual between original and dehazed images can serve as an explanatory variable related to haze effects for predicting PM levels. Motivated by this, we set a model of the relationship between the variables of haze effects and PM levels. Using components stored in estimated haze effects $\hat{\mathbf{H}}$ (in Figure 3), we can estimate PM concentration levels.

3.2. Temporal Prior. Since the PM consists of infinitesimal particles floating in the air, they can be characterized by a transport phenomenon such as flow motions in video. Although we cannot directly identify flow motion of PM by tracking all particles, we can indirectly discover its effects through variability in the midst of multiple image sequences in video. As introduced by Kim et al. [37], differences between the original consecutive frames in both cases (safe $\leq 80 \mu\text{g}/\text{m}^3$ and harmful $> 80 \mu\text{g}/\text{m}^3$) are obviously distinctive. This previous work reinforces the plausibility of hypothesis that the image distinctions across images show PM concentration levels are considerably significant. This is motivated by the fact that the moving objects happen to create subtle changes. Stepping beyond the previous work, inspired by prior in Bayesian, we impose the fluid flow model proposed by Xie et al. [38] in order to additionally accommodate feature variations:

$$\mathbf{F}_t = \mathbf{T}(\mathbf{F}_{t-1}, \mathbf{P}_{t-1}), \quad (5)$$

where t is the time domain and \mathbf{F} , \mathbf{T} , and \mathbf{P} denote the fluid, the transport operator (a.k.a. advection or warp operator), and the temporal prior, respectively. Note that $\mathbf{p}_t \in \mathbb{R}^2$, the element of $\mathbf{P}_t \in \mathbb{R}^3$, is the vector, which has the velocity and flow direction field. For brevity, we replace the existing physical prior \mathbf{P}_{t-1} , where \mathbf{P}_{t-1} denotes the average over a period of time prior to t . More precisely, Figure 4 describes how the proposed temporal prior is made up based on sequential frames during the prior length, each of pixels calculating the average values of RGB across image frames. Taken together, we combine the original RGB (3 channels) and its corresponding temporal priors (3 channels) into augmented inputs for a total of six channels (see the first half of the network in Figure 2). When exploiting the priors almost consistent across image sequences, the network is expected to be superior in producing consistent haze effects, as compared to the model with no temporal prior.

4. Numerical Experiments

4.1. Datasets. In this section, we describe haze video datasets to apply. One of the challenges in creating real video datasets is that both consecutive haze frames and the corresponding haze-free frames are supposed to be perfectly matched up. To circumvent a little, artificial haze video datasets based on existing videos can be used to train dehazing networks [39–41]. Yet, synthetic haze effects hardly accommodate natural flow motions that disperse light, and thereby, this wrongly distorts pixel values. Instead, we collected datasets of various environments including both indoor and outdoor areas (refer to the sample haze images in Supplementary Material (available here)). Here, one note of caution is that since indoor environments are not directly exposed to the outside atmosphere, in the case of indoor, it is difficult to collect video clips with high PM levels. Therefore, we had to open all the windows when PM was high or sometimes directly generate PM by burning incense or spraying potassium chloride. The thumbnail images are shown in Table 2. The video clips were recorded with the *Raspberry Pi Camera Module V2* (Raspberry Pi Foundation; <http://www.raspberrypi.org/>) employing a low-cost CMOS sensor. In the stage of building datasets, we first mount the pi cameras on the tripods for image stabilization. While recording video clips, we concurrently measure PM levels using the *Aerocet 831 Handheld Particle Counter* (Met One Instruments; <http://www.metone.com/>), a high-precision device. With the help of the *Aerocet*, we collected the haze-free images under atmosphere where PM levels are less than or equal to 15 as the target images for training the network. For these data, we finally applied the optical flow [42] to the captured video clips for removing any spatial variances. All the datasets and the codes are available at the author’s website (<http://www.hifiai.pe.kr/>).

4.2. Results. To measure PM levels, the estimated haze effects per frame (matrix) are converted to statistic (scalar) such as mean, entropy, and variance [20, 43]. To assess

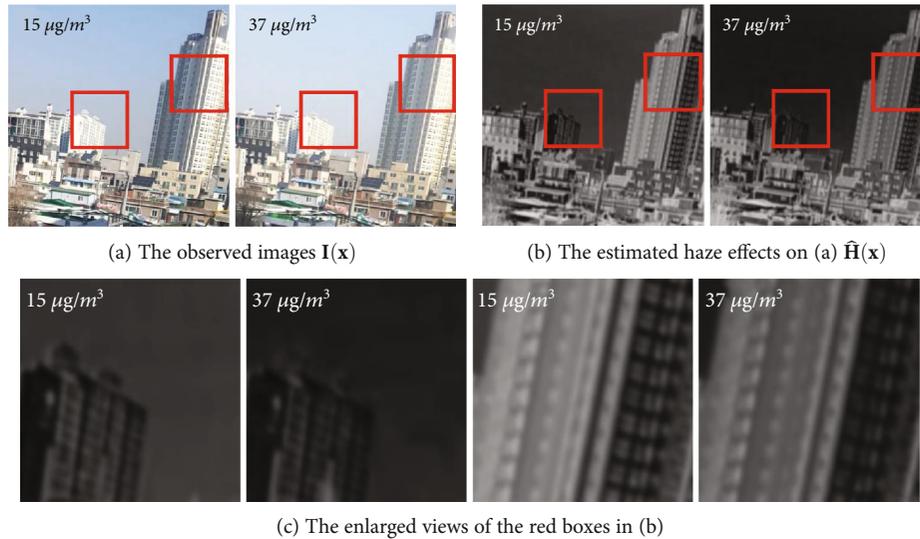


FIGURE 3: The differences between the estimated haze effects varied by PM concentration level ($15 \mu\text{g}/\text{m}^3$ and $37 \mu\text{g}/\text{m}^3$). The estimated haze effect images in RGB channels were converted to grayscale image for a clear comparison.

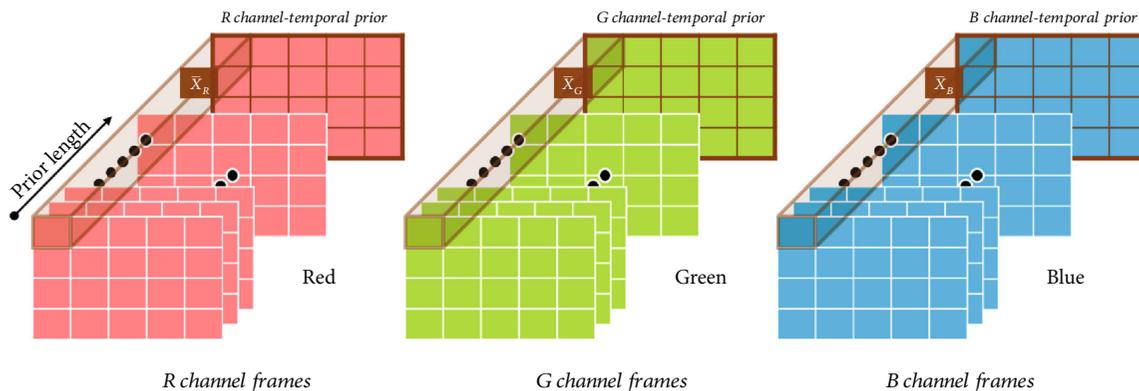


FIGURE 4: The illustration of temporal prior using RGB channels.

predictive power, we compare the true hazy effects with predicted hazy effects. To do so, we fit the three regression-type models: (1) random forest regression (RFR), (2) support vector regression (SVR) with the radial basis function (rbf) kernel, and (3) multilayer perceptron regression (MPR). The metric to evaluate prediction accuracy is as follows:

$$\text{acc} = \frac{1 - (|AQI_{\text{ground}} - AQI_{\text{predicted}}|)}{QI_{\text{ground}}}, \quad (6)$$

where AQI_{ground} and $AQI_{\text{predicted}}$ refer to statistic (e.g., mean, entropy, and variance) of $\mathbf{H}(\mathbf{x})$ from ground truth and the proposed model. Table 3 summarizes the accuracy results for the test sets, and all indoor and outdoor scenarios are equally assigned.

4.2.1. Indoor and Outdoor Environment. We collect over 2,000 video clips from the indoor office and corridors at the

Konkuk University over several months. Here, the proposed model presents the best performance of accuracy with 86.72% when adopting MLPR with prior and the entropy benchmark. Especially notable is that accuracy tends to increase when applying the temporal prior across all scenarios. In this sense, it is confirmed that temporal priors can provide the network with additional haze-related information. In addition, indoor environments are believed to be less sensitive to environmental factors, and thus, it is understandable that indoor experiments outperform outdoor as a whole. For outdoor experiments, we select two locations in South Korea populated with people in the midst of the residential areas and the building complex. For several months, we collect more than 3,000 video clips of each location. For reliability, we make sure that outdoor data consist of a widespread range of PM levels. In Table 3, SVR with entropy presents 72.16% and 82.45% for nonprior and prior, respectively. Interestingly, priors in outdoor data allow considerable accuracy

TABLE 2: Brief descriptions of selected regions and experiment chambers.

Category	Thumbnail	# of video clips
Indoor environments		2,000 for each site
Outdoor environments		3,000 for each site
Experimental chamber		1,000 for each experimental condition

gains compared to indoor scenarios. The results implicate that prior information can be effective especially in outdoor environments (e.g., windiness) in the midst of particulate flows.

4.2.2. Experimental Chamber. In simulations, it is essential to assess diverse environmental conditions. To this end, we especially design the experimental chamber to implement diverse conditions of interest. The experiments were carried out regarding four factors including wind, temperature, humidity, and illuminance. In the course of the experiments, the other confounding factors are well adjusted in the experiment chamber. We gather approximately 1,000 video clips across all experiment conditions. In Table 4, the results show that the MLPR with entropy consistently present high accuracy over 80% for almost all scenarios. Therefore, it is clearly confirmed that this prior-based model serves to adequately control environmental confounding factors that possibly intervene in haze effects.

5. Discussion

Undoubtedly, the latest AI has mainly focused on vision-based techniques (e.g., RGB and Lidar). Nonetheless, in the AI domain, infinitesimal materials still remain underestimated due to its invisible nature. In this regard, vision-based PM measurements are featured with many advantages of flexibility and accessibility in the view of real-time air quality monitoring and extension to spatial scales. Given our finding, even with a low-cost optical sensor, the proposed method can offer further benefits in business and cost savings in practical aspects. Moreover, this also serves as a predictive

TABLE 3: The accuracy results of the regression models for various environments. The best results pertaining to each condition are underlined.

	Nonprior			Prior		
	RFR	SVR	MLPR	RFR	SVR	MLPR
Indoor						
<i>Mean</i>	0.6903	0.5872	0.7546	0.7299	0.6139	0.7953
<i>Variance</i>	0.6645	0.8101	0.8152	0.6682	0.8239	0.8258
<i>Entropy</i>	0.6949	<u>0.8247</u>	0.8240	0.7054	0.8388	<u>0.8672</u>
Outdoor						
<i>Mean</i>	0.6228	0.4599	0.5882	0.6917	0.5197	0.6969
<i>Variance</i>	0.6157	0.6843	0.6891	0.7005	0.7760	0.7753
<i>Entropy</i>	0.6193	<u>0.7216</u>	0.7135	0.7127	<u>0.8245</u>	0.8132

model using the deep cascaded CNN and temporal prior in a methodological aspect. Compared to existing vision-based predictive models, the proposed model stretches into accommodating additional temporal prior features among image sequences, aiming at improving predictive power in the virtue of data augmentation. With various simulation designs (e.g., real data and experimental chamber data), we confirm that the proposed models are superior to the traditional models without temporal priors, showing outstanding predictive power. Further improvements can be made by exploiting the optimal length of frames in the context of optimizing predictive power. This effort can facilitate to export it to a gauging device and promote practical utility, since all of vision-based models strongly depend on well-controlled

TABLE 4: Experimental chamber study. The accuracy results of the regression models for various environmental conditions. The best results pertaining to each condition are underlined.

Nonprior	Normal state			Windiness			Prior			Normal state			Windiness		
	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR
Mean	0.7586	0.6412	0.7431	0.6403	0.5986	0.6653	0.7620	0.6441	0.7899	0.6951	0.4670	0.7235	0.6951	0.4670	0.7235
Variance	0.7116	<u>0.7805</u>	0.7470	0.6319	0.6951	0.7249	0.6728	0.7995	0.8093	0.6357	0.6428	0.7764	0.6357	0.6428	0.7764
Entropy	0.6511	0.7523	0.7528	0.7074	<u>0.7780</u>	0.6057	0.6934	0.8674	<u>0.8699</u>	0.7186	<u>0.8428</u>	0.8101	0.7186	<u>0.8428</u>	0.8101

Nonprior	Temperature (low) (20°C)			Temperature (high) (40°C)			Prior			Temperature (low) (20°C)			Temperature (high) (40°C)		
	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR
Mean	0.6906	0.5189	0.7254	0.6105	0.5143	0.6856	0.7154	0.5162	0.8123	0.7636	0.5810	0.8094	0.7636	0.5810	0.8094
Variance	0.7101	0.7567	0.7674	0.7083	0.7410	0.7528	0.6510	0.7854	<u>0.8518</u>	0.6883	0.7681	0.8082	0.6883	0.7681	0.8082
Entropy	0.7517	<u>0.7781</u>	0.7056	0.6505	<u>0.7696</u>	0.7387	0.7215	0.8304	0.8319	0.7271	0.8330	<u>0.8334</u>	0.7271	0.8330	<u>0.8334</u>

Nonprior	Humidity (low) (25%)			Humidity (high) (50%)			Prior			Humidity (low) (25%)			Humidity (high) (50%)		
	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR
Mean	0.6910	0.6397	0.6666	0.6593	0.6317	0.7234	0.7319	0.5523	0.7854	0.7462	0.5699	0.7734	0.7462	0.5699	0.7734
Variance	0.7218	0.7581	0.6977	0.6745	0.6612	0.6715	0.6821	0.8149	0.7724	0.6770	0.7675	0.7379	0.6770	0.7675	0.7379
Entropy	0.7203	<u>0.7827</u>	0.6660	0.6679	<u>0.7674</u>	0.6869	0.7466	0.8054	<u>0.8524</u>	0.7451	0.8174	<u>0.8468</u>	0.7451	0.8174	<u>0.8468</u>

Nonprior	Illuminance (low) (100 lx)			Illuminance (high) (300 lx)			Prior			Illuminance (low) (100 lx)			Illuminance (high) (300 lx)		
	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR	RFR	SVR	MLPR
Mean	0.6659	0.5286	0.7038	0.6562	0.5634	0.7140	0.7131	0.5506	0.6905	0.6672	0.5185	0.7522	0.6672	0.5185	0.7522
Variance	0.6433	0.7024	0.6046	0.6450	0.7513	0.6685	0.6241	0.7861	0.7757	0.7384	0.7135	0.8104	0.7384	0.7135	0.8104
Entropy	0.6588	<u>0.7821</u>	0.6280	0.6615	<u>0.7764</u>	0.6800	0.6825	0.8018	<u>0.8031</u>	0.6926	0.8355	<u>0.8764</u>	0.6926	0.8355	<u>0.8764</u>

radiance, which considerably discourages vision-based measurement techniques at times. To address this issue, we plan on developing alternative prediction models that account for particulate-related features only. We leave these topics for future research.

Data Availability

All the datasets and the codes are available at the author's website (<http://www.hifai.pe.kr/>).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Konkuk University Researcher Fund in 2019, Konkuk University Researcher Fund in 2020, and the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2020R1C1C1A01005229).

Supplementary Materials

Fig. S1: the sample haze images. (a) Konkuk University office. (b) The residential area. (c) The experimental chamber. (*Supplementary Materials*)

References

- [1] K. H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," *Environment International*, vol. 74, pp. 136–143, 2015.
- [2] February 2020, http://www.nbrienvic.nic.in/Database/1_2463.aspx.
- [3] Y. Wang, "The analysis of the impacts of energy consumption on environment and public health in China," *Energy*, vol. 35, no. 11, pp. 4473–4479, 2010.
- [4] J. D. Sacks, L. W. Stanek, T. J. Luben et al., "Particulate matter-induced health effects: who is susceptible?," *Environmental Health Perspectives*, vol. 119, no. 4, pp. 446–454, 2011.
- [5] A. Mukherjee and M. Agrawal, "World air particulate matter: sources, distribution and health effects," *Environmental Chemistry Letters*, vol. 15, no. 2, pp. 283–309, 2017.
- [6] Q. Zhang, Y. Niu, Y. Xia et al., "The acute effects of fine particulate matter constituents on circulating inflammatory biomarkers in healthy adults," *Science of the Total Environment*, vol. 707, article 135989, 2020.
- [7] O. Raaschou-Nielsen, Z. J. Andersen, R. Beelen et al., "Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE)," *The Lancet Oncology*, vol. 14, no. 9, pp. 813–822, 2013.
- [8] H. C. Kim, S. Kim, B.-U. Kim et al., "Recent increase of surface particulate matter concentrations in the Seoul Metropolitan Area, Korea," *Scientific Reports*, vol. 7, no. 1, pp. 1–7, 2017.

- [9] OECD, *OECD Economic Surveys: Korea 2018*, OECD Publishing, 2018.
- [10] 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*, pp. 1207–1215, Paris, France, France, 2019.
- [11] Y. Cheng, X. Li, Z. Li et al., "AirCloud: a cloud-based air-quality monitoring system for everyone," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems - SenSys '14*, pp. 251–265, Memphis, Tennessee, USA, 2014.
- [12] March 2020, <https://www.thermofisher.com/order/catalog/product/TEOM1405>.
- [13] March 2020, <https://metone.com/products/aerocet-831-handheld-particle-counter>.
- [14] March 2020, <http://www.dylospromod.com/dcproairqumo.html>.
- [15] March 2020, <https://www.airkorea.or.kr/web/stationInfo?pMENU=93>.
- [16] Y. Zheng, F. Liu, and H. P. Hsieh, "U-air: when urban air quality inference meets big data," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '13*, pp. 1436–1444, Chicago, Illinois, USA, 2013.
- [17] Y. Gao, W. Dong, K. Guo et al., "Mosaic: a low-cost mobile sensing system for urban air quality monitoring," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pp. 1–9, San Francisco, CA, USA, 2016.
- [18] J. Li, Q. Fu, J. Huo et al., "Tethered balloon-based black carbon profiles within the lower troposphere of Shanghai in the 2013 East China smog," *Atmospheric Environment*, vol. 123, pp. 327–338, 2015.
- [19] K. Weber, G. Heweling, C. Fischer, and M. Lange, "The use of an octocopter UAV for the determination of air pollutants—a case study of the traffic induced pollution plume around a river bridge in Duesseldorf, Germany," *International Journal of Education and Learning Systems*, vol. 2, 2017.
- [20] C. Liu, F. Tsow, Y. Zou, and N. Tao, "Particle pollution estimation based on image analysis," *PLoS One*, vol. 11, no. 2, article e0145955, 2016.
- [21] S. Li, T. Xi, Y. Tian, and W. Wang, "Inferring fine-grained PM_{2.5} with Bayesian based kernel method for crowdsourcing system," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, Singapore, Singapore, 2017.
- [22] Z. Pan, H. Yu, C. Miao, and C. Leung, "Crowdsensing air quality with camera-enabled mobile devices," in *Twenty-Ninth IAAI Conference*, San Francisco, California, USA, 2017.
- [23] H. Wang, X. Yuan, X. Wang, Y. Zhang, and Q. Dai, "Real-time air quality estimation based on color image processing," in *2014 IEEE Visual Communications and Image Processing Conference*, pp. 326–329, Valletta, Malta, 2014.
- [24] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, pp. 228–242, 2007.
- [25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [26] 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, 2011.
- [27] C. Li, J. Guo, F. Porikli, H. Fu, and Y. Pang, "A cascaded convolutional neural network for single image dehazing," *IEEE Access*, vol. 6, pp. 24877–24887, 2018.
- [28] V. Natarajan, "Enhanced single image uniform and heterogeneous fog removal using guided filter," in *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, S. Dash, K. Vijayakumar, B. Panigrahi, and S. Das, Eds., vol. 517 of *Advances in Intelligent Systems and Computing*, pp. 453–463, Springer, Singapore, 2017.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [30] S. Han, H. Mao, and W. J. Dally, "Deep compression: compressing deep neural networks with pruning, trained quantization and Huffman coding," 2015, <https://arxiv.org/abs/1510.00149>.
- [31] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in *Advances in Neural Information Processing Systems*, pp. 1135–1143, Curran Associates, Inc., 2015.
- [32] J. Yamanaka, S. Kuwashima, and T. Kurita, "Fast and accurate image super resolution by deep CNN with skip connection and network in network," in *Neural Information Processing*, D. Liu, S. Xie, Y. Li, D. Zhao, and E. S. El-Alfy, Eds., vol. 10635 of *ICONIP 2017. Lecture Notes in Computer Science*, pp. 217–225, Springer, Cham, 2017.
- [33] D. M. Strong, P. Blomgren, and T. F. Chan, "Spatially adaptive local-feature-driven total variation minimizing image restoration," in *Statistical and Stochastic Methods in Image Processing II*, vol. 3167, pp. 222–233, San Diego, CA, USA, 1997.
- [34] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, <https://arxiv.org/abs/1312.4400>.
- [35] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, 2015.
- [36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097–1105, Curran Associates Inc., 2012.
- [37] S. Kim, S. Jung, S. Yang et al., "Vision-based deep Q-learning network models to predict particulate matter concentration levels using temporal digital image data," *Journal of Sensors*, vol. 2019, 10 pages, 2019.
- [38] Y. Xie, E. Franz, M. Chu, and N. Thuerey, "tempoGAN: a temporally coherent, volumetric gan for super-resolution fluid flow," *ACM Transactions on Graphics*, vol. 37, no. 4, pp. 1–15, 2018.
- [39] B. Cai, X. Xu, and D. Tao, "Real-time video dehazing based on spatiotemporal mrf," in *Advances in Multimedia Information Processing - PCM 2016. PCM 2016*, E. Chen, Y. Gong, and Y. Tse, Eds., vol. 9917 of *Lecture Notes in Computer Science*, pp. 315–325, Springer, Cham, 2016.
- [40] W. Ren and X. Cao, "Deep video dehazing," in *Advances in Multimedia Information Processing - PCM 2017. PCM 2017*, B. Zeng, Q. Huang, A. Saddik, H. Li, S. Jiang, and X. Fan,

Eds., vol. 10735 of Lecture Notes in Computer Science, pp. 14–24, Springer, Cham, 2017.

- [41] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “End-to-end united video dehazing and detection,” in *Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, Louisiana, USA, 2018.
- [42] D. Sun, S. Roth, and M. J. Black, “Secrets of optical flow estimation and their principles,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2432–2439, San Francisco, CA, USA, 2010.
- [43] W. Yuchi, E. Gombojav, B. Boldbaatar et al., “Evaluation of random forest regression and multiple linear regression for predicting indoor fine particulate matter concentrations in a highly polluted city,” *Environmental Pollution*, vol. 245, pp. 746–753, 2019.