Environment International 144 (2020) 106060

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

New interpretable deep learning model to monitor real-time $PM_{2.5}$ concentrations from satellite data

Xing Yan[a], Zhou Zang[a], Nana Luo[a,b], Yize Jiang[a], Zhanqing Li[c,*]

[a] State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China
[b] Department of Geography, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182-4493, USA
[c] Department of Atmospheric and Oceanic Sciences and ESSIC, University of Maryland, College Park, MD, USA

ARTICLE INFO

Handling Editor: Xavier Querol

Keywords:
Deep learning
Satellite
$PM_{2.5}$
Himawari-8

ABSTRACT

Particulate matter with a mass concentration of particles with a diameter less than 2.5 μm ($PM_{2.5}$) is a key air quality parameter. A real-time knowledge of $PM_{2.5}$ is highly valuable for lowering the risk of detrimental impacts on human health. To achieve this goal, we developed a new deep learning model-EntityDenseNet to retrieve ground-level $PM_{2.5}$ concentrations from Himawari-8, a geostationary satellite providing high temporal resolution data. In contrast to the traditional machine learning model, the new model has the capability to automatically extract $PM_{2.5}$ spatio-temporal characteristics. Validation across mainland China demonstrates that hourly, daily and monthly $PM_{2.5}$ retrievals contain the root-mean-square errors of 26.85, 25.3, and 15.34 μg/m^3, respectively. In addition to a higher accuracy achievement when compared with various machine learning inversion methods (backpropagation neural network, extreme gradient boosting, light gradient boosting machine, and random forest), EntityDenseNet can "peek inside the black box" to extract the spatio-temporal features of $PM_{2.5}$. This model can show, for example, that $PM_{2.5}$ levels in the coastal city of Tianjin were more influenced by air from Hebei than Beijing. Further, EntityDenseNet can still extract the seasonal characteristics that demonstrate that $PM_{2.5}$ is more closely related within three month groups over mainland China: (1) December, January and February, (2) March, April and May, (3) July, August and September, even without meteorological information. EntityDenseNet has the ability to obtain high temporal resolution satellite-based $PM_{2.5}$ data over China in real-time. This could act as an important tool to improve our understanding of $PM_{2.5}$ spatio-temporal features.

1. Introduction

Particulate matter with an aerodynamic diameter less than 2.5 μm ($PM_{2.5}$, i.e., a fine-mode aerosol) is an environmental parameter of great concern worldwide (Apte et al., 2015; Yan et al., 2014). The ground-level $PM_{2.5}$ concentrations have been monitored around the world (Liang et al., 2020; van Donkelaar et al., 2015; Yang et al., 2018), but spatial coverage is lacking and inhomogeneous, making it difficult to capture its variability and patterns (Han et al., 2015). To overcome this limitation, satellite-based remote sensing of aerosols (aerosol optical depth, AOD) has been widely used to estimate $PM_{2.5}$ (van Donkelaar et al., 2006; Liu et al., 2009; Xu et al., 2018; Lin et al., 2018; Hu et al., 2019; Sayer et al., 2014).

Various remote sensing approaches have been developed to retrieve $PM_{2.5}$ from satellite-derived AOD (Guo et al., 2014; Gupta et al., 2006; Hu et al., 2017; Koelemeijer et al., 2006; Kokhanovsky et al., 2009; Ma et al., 2014; Sorek-Hamer et al., 2015; Zhang and Li, 2015). However, inherent limitations still exist. The AOD is a columnar quantity encompassing aerosols of all sizes. Fine mode AOD (fAOD) is correlated more closely with $PM_{2.5}$ and is a better predictor for estimating surface $PM_{2.5}$ concentration (Zhang and Li, 2013; Yan et al., 2017). However, the fine mode fraction (FMF, fAOD = AOD × FMF) varies largely over land and is subject to great uncertainties during satellite retrievals (Yan et al., 2019), and thus is not generated in the current version of C6 MODIS global scale aerosol products (Levy et al., 2013). Further, in AOD-$PM_{2.5}$ modeling, rich spectral information contained in satellite measurements is rarely used. In addition to the satellite visible band, the short-wave infrared (Shen et al., 2018) and thermal infrared (Amanollahi et al., 2013) bands can also provide unique information required for estimating surface particulate matter concentration. Unfortunately, the application of full spectral information from satellite for $PM_{2.5}$ retrieval is still highly limited.

* Corresponding author.
E-mail address: zli@atmos.umd.edu (Z. Li).

https://doi.org/10.1016/j.envint.2020.106060
Received 18 June 2020; Received in revised form 7 August 2020; Accepted 14 August 2020

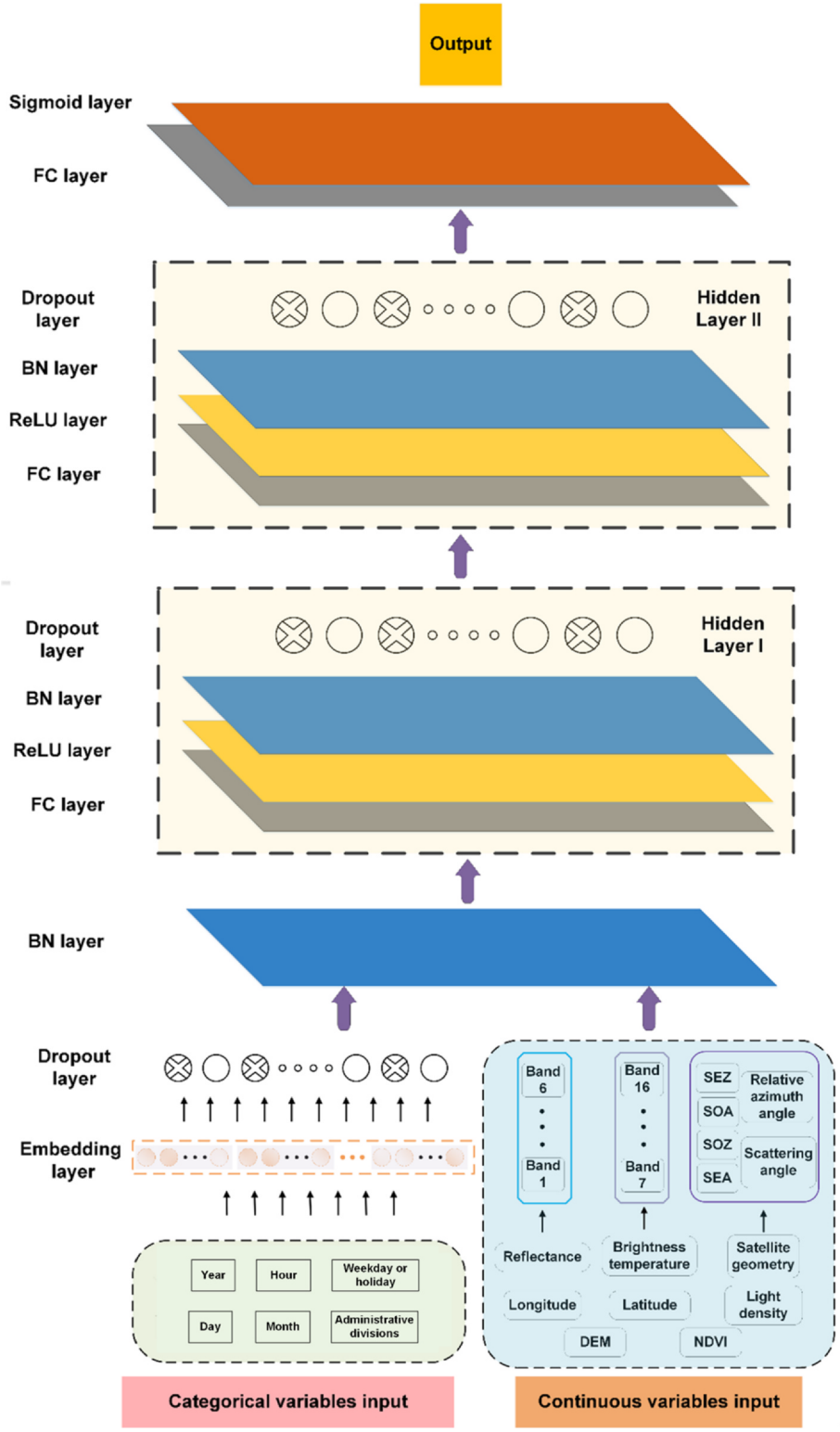

Fig. 1. Schematic diagram of the proposed EntityDenseNet in this study.

Polar-orbiting satellites only provide one measurement daily, during daytime hours, and this makes it incapable of monitoring temporal evolution of air quality. The Himawari-8, a new-generation Japanese meteorological satellite launched on 7 October 2014 and first released data on 7 July 2015, provides regional AOD data at a 10-min interval, but no real-time $PM_{2.5}$ data. Hoff and Christopher (2009) reviewed over 200 studies and concluded that the precision of $PM_{2.5}$ from AOD was ± 30% (maximum) and estimates were only achieved with additional input parameters (e.g., meteorological and atmospheric profile information). However, obtaining meteorological and atmospheric profile data at equal satellite spatial resolution (e.g., 5 km or 1 km) and coverage by model simulation is highly time consuming, and this hinders real-time $PM_{2.5}$ retrieval. Hence, the ability for satellite to still achieve ± 30% precision to monitor $PM_{2.5}$ without meteorological or atmospheric profile information is under question.

In recent years, machine learning methods have increased in

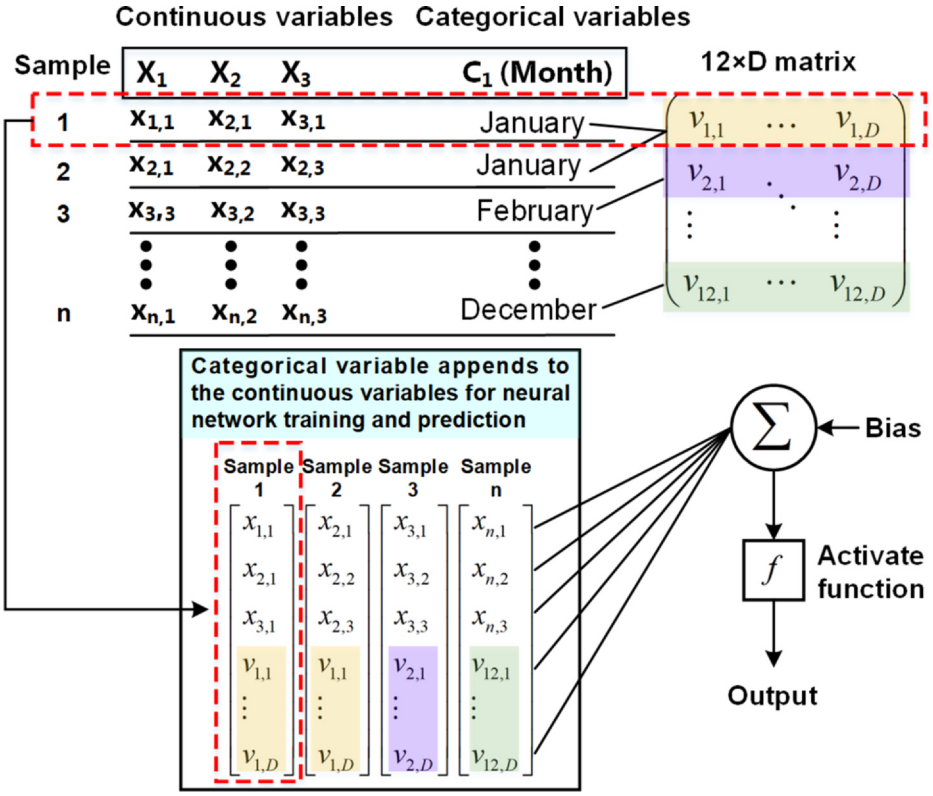

Fig. 2. Diagram of the categorical variable "Month" appended to the continuous variables for the EntityDenseNet training and prediction.

popularity as a method for estimating $PM_{2.5}$ concentrations using satellite data (Liu et al., 2017; Mao et al., 2017). However, classical machine learning methods possess intrinsic limitations, which makes it challenging for satellite-based $PM_{2.5}$ retrieval. For example, the neural network approach cannot directly use categorical data (Guo and Berkhahn, 2016). The neural network requires all input variables and output variables to be numeric (Yan et al., 2015), thus, the One-Hot encoding is typically used for converting the categorical variable which is then input into neural network training and prediction (Chren, 1998; Liu et al., 2002). However, there are two key shortcomings to One-Hot encoding: (1) it often requires an unrealistic level of computational resources, and; (2) it treats different categorical variables absolutely independently of each other and often ignores the informative relationships between them (Guo and Berkhahn, 2016). Therefore, as deep learning is usually a neural network-based model, one of the key issues to overcome was determining how to handle and learn the information using categorical variables. Finally, interpreting the spatio-temporal features from deep learning neural networks also remains challenging. Although Reichstein et al. (2019) indicated that deep machine learning models may act as a promising tool to extract spatial–temporal features from the data, the processes required to open and interpret this "black box" model are difficult.

In this study, we propose a new interpretable deep learning model called EntityDenseNet to estimate real-time ground-level $PM_{2.5}$ from Himawari-8 satellite data. The EntityDenseNet is able to extract $PM_{2.5}$ spatio-temporal characteristics automatically during the training process. In contrast to the previous models, the EntityDenseNet uses full spectral information to directly retrieve $PM_{2.5}$ concentrations. A comprehensive discussion of this method and comparisons with BPNN, extreme gradient boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and RF is also presented.

2. Data and methods

2.1. Himawari-8 data

The Himarwari-8 reflectance (Bands 1 to 6) and brightness temperature (Bands 7 to 16) data of a spatial resolution of 5 km were extracted from the L1 Gridded Data at 10 min intervals during daytime (UTC 1:00–6:00) from January 2016 to June 2019. The cloud mask was based on Ishida and Nakajima (2009). The Top-Of-Atmosphere (TOA) reflectance used as the input data for the proposed network is calculated from the albedo reflectance data:

$$\rho^* = \frac{albedo}{\cos(\theta_0)} \quad (1)$$

where ρ^* is the TOA reflectance and θ_0 is the solar zenith angle.

2.2. $PM_{2.5}$, elevation and light density data

Hourly $PM_{2.5}$ concentrations were collected for the same period from 1434 monitoring stations across mainland China. The Tapered Element Oscillating Microbalance (TEOM) was used to measure $PM_{2.5}$ concentration at each site. You et al. (2016) showed that the uncertainty in $PM_{2.5}$ measured daily is $\pm$ 1.5 μg/m^3. In this study, to test the ability of the proposed model in dealing with abnormal values, no extreme $PM_{2.5}$ values (e.g., $PM_{2.5}$ > 800 μg/m^3) are excluded.

The digital elevation model (DEM) data of 30 m resolution was obtained from the Shuttle Radar Topography Mission (http://srtm.csi.cgiar.org). Light density data were provided by the Earth Observation Group of the NOAA National Centers for Environmental Information (NCEI). We downloaded the annual 2016 Version 1 Nighttime VIIRS Day/Night (VIIRS Cloud Mask - Outlier Removed - Nighttime Lights) for this study (https://ngdc.noaa.gov/eog/viirs/).

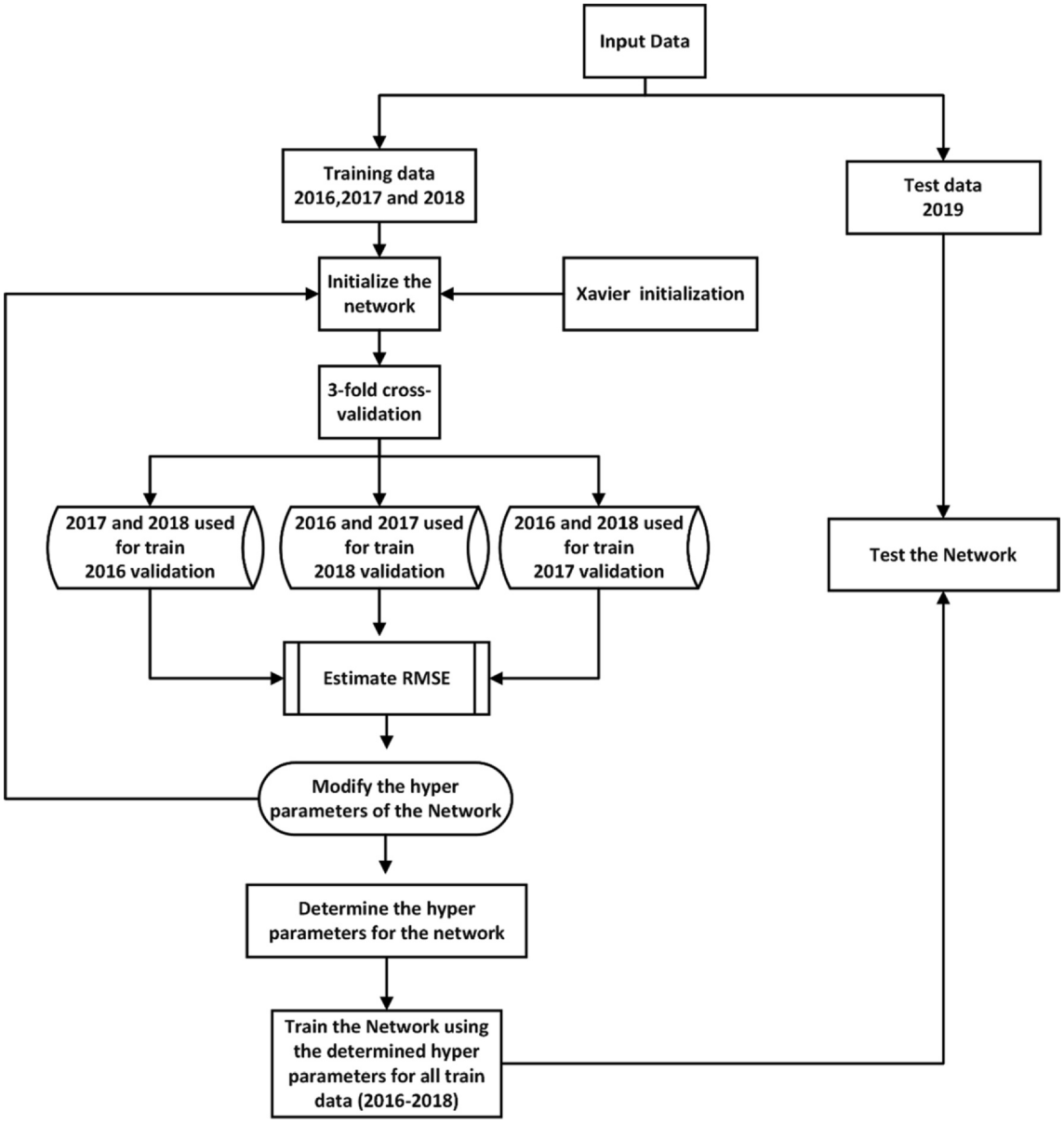

Fig. 3. Flowchart describing the entire training process of EntityDenseNet.

2.3. EntityDenseNet

The structure of EntityDenseNet method is presented in Fig. 1. The input data is first separated into two data types: categorical variables and continuous variables. The categorical variables include year, month, date, hour, China administrative divisions and day type (weekday or holiday). The continuous variables consist of TOA reflectance from Himawari-8 bands 1 to 6, brightness temperature data from Himawari-8 bands 7 to 16, satellite zenith angle (SEZ), solar zenith angle (SOZ), satellite azimuthal angle (SEA), solar azimuthal angle (SOA), relative azimuth angle, scattering angle, longitude, latitude, light density, Digital Elevation Model (DEM), and Normalized Vegetation Index (NDVI). The details of calculating scattering angle and NDVI are shown in the Supplementary material.

To process the categorical variables, rather than the One-Hot encoding method, we used the Entity Embeddings method (Guo and Berkhahn, 2016). Guo and Berkhahn (2016) demonstrated that this method not only speeds up neural networks compared with one-hot encoding, but also helps the neural network to learn the intrinsic relationship between the categorical variables. In the proposed EntityDenseNet, we introduced the embedding layer to implement the Entity Embeddings method, and the purpose of this layer is to map the discrete categorical variable to the vector of continuous numbers. In the embedding layer, a random matrix was first initialized as m × D:

$$m \times Drandommatrix \rightarrow \begin{pmatrix} v_{1,1} & \cdots & v_{1,D} \\ \vdots & \ddots & \vdots \\ v_{m,1} & \cdots & v_{m,D} \end{pmatrix}$$

where m is the number of unique levels of a categorical variable (for example, if we want to map month categorical variable Jan-Dec to embedding vectors, the m is set to 12) and D is a hyper parameter that can be between 1 and m − 1. It should be noted that we do not have to specify the m × D matrix values manually as they are trainable parameters, in other words, matrix values learned by the model during training. Thus, first, the m × D matrix is randomly initialized and set as parameters to EntityDenseNet model. Then, the m × D matrix values are gradually adjusted via backpropagation during the training process, and finally it is determined when the model has been trained. In Fig. 2, we use the categorical variable "Month" as an example to show how the embedding layer works in this EntityDenseNet model. A 12 × D matrix

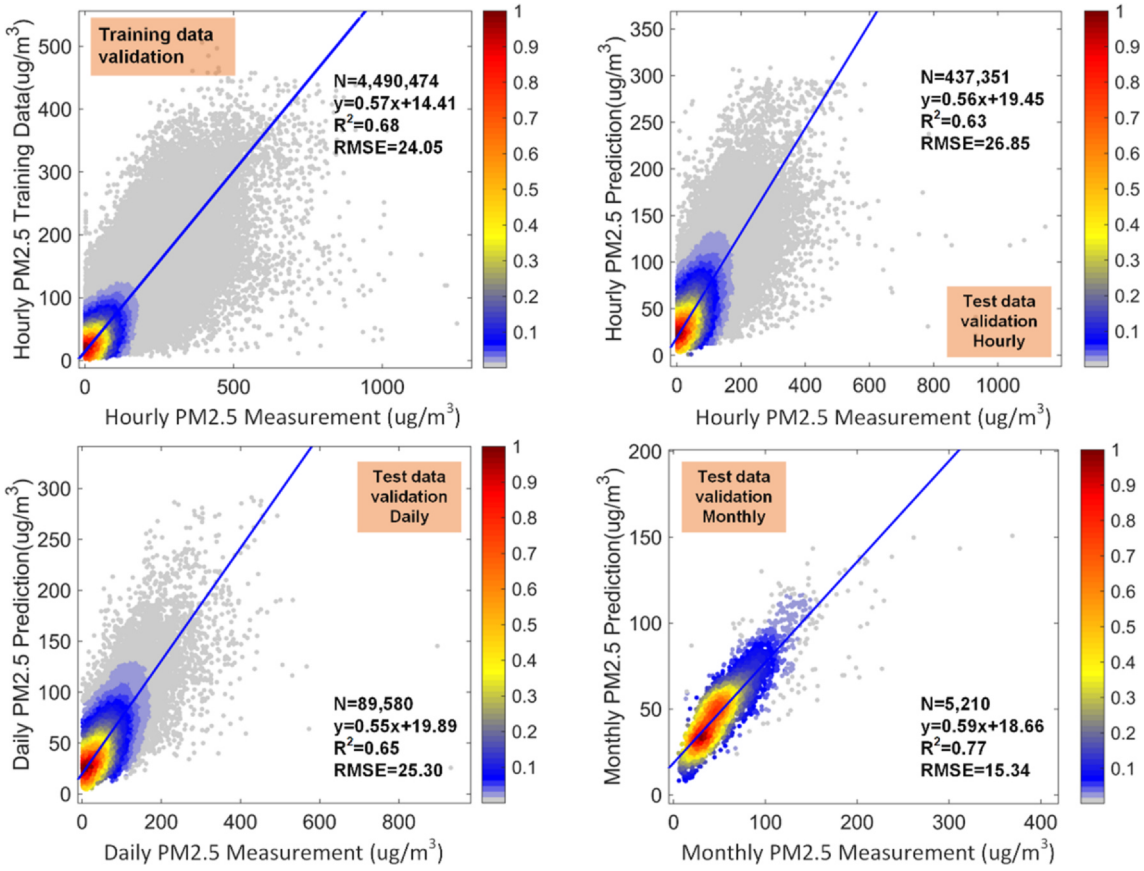

Fig. 4. Scatterplot showing the performance of EntityDenseNet on trained (2016–2018) and test (2019,1.1–6.1) data sets.

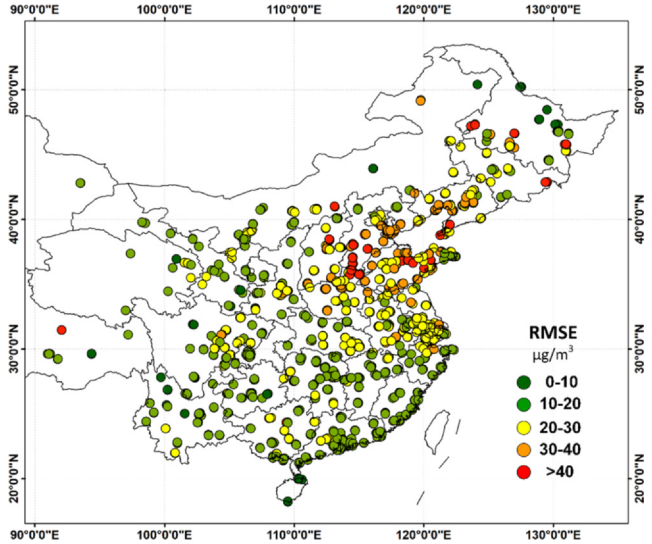

Fig. 5. RMSE values for the EntityDenseNet estimates of $PM_{2.5}$ in 1434 monitoring sites.

is randomly initialized first, and each row (1 × D vector) of this matrix corresponds to specific month information. This 1 × D vector is appended to the continuous variables for neural network training and prediction as shown in Fig. 2. The values of hyperparameter D and v in the 12 × D matrix is updated and finally determined during the network training phase.

Each continuous variable $x_i$is first normalized by the Z-score method (is shown in the Supplementary material) before being inputted into the EntityDenseNet (Ben Khalifa et al., 2013). As shown in Fig. 1, the merged categorical and continuous variables are input into a batch normalization (BN) layer. The purpose of the BN layer is to ensure that the data in each layer is distributed equally, which greatly accelerates the learning speed of neural networks (Ioffe and Szegedy, 2015). Each of the two hidden layers in EntityDenseNet includes one fully connected layer, one rectified linear unit (ReLU) layer, one BN layer, and one dropout layer (Yan et al., 2020). We used the ReLU as the activation function in this study because it can overcome the problems of saturation and vanishing gradients (Nair and Hinton, 2010). To prevent overfitting, we introduced a dropout layer to each hidden layer. The dropout method has been demonstrated to significantly reduce overfitting and improve the performance of neural networks (Srivastava et al., 2014). The feed-forward operation in the hidden layer of EntityDenseNet can be described as:

$$G^m = r^m \times y^m \tag{4}$$

$$x_i^{m+1} = W_i^{m+1} G^m + b_i^{m+1} \tag{5}$$

$$R_i^{m+1} = BN\left[f\left(x_i^{m+1}\right)\right] \tag{6}$$

where m is the index of the hidden layer, y^m is the vector of outputs from layer m, G^m is the thinned vector of y^m by the dropout layer. r^m is a vector of independent Bernoulli random variables, each of which has probability p of being 1 and probability 1-p of being 0. x^m is the vector of inputs into layer m; W^m and b^m are the weights and biases of layer m, respectively. R_i^{m+1} is the output from hidden layer $m + 1$ at the neuron node i. The $f()$ function is the ReLU activation function (Supplementary), $BN[]$ is the BN function, the details of which are discussed in Ioffe and Szegedy (2015).

2.4. Network training

Fig. 3 is a flowchart describing the entire training process of

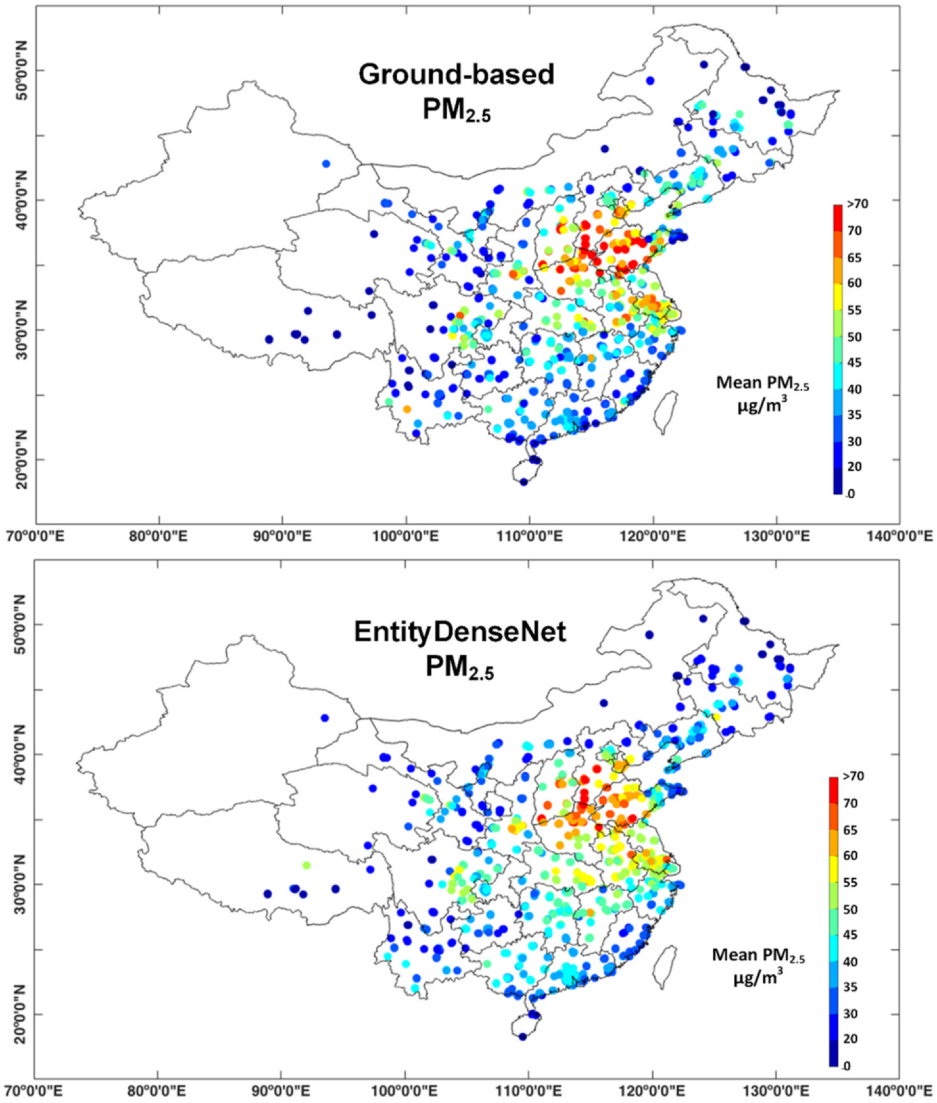

Fig. 6. Mean $PM_{2.5}$ concentration during the test data period from 1 January 2019 to 1 June 2019 from ground-based measurements and EntityDenseNet retrieval.

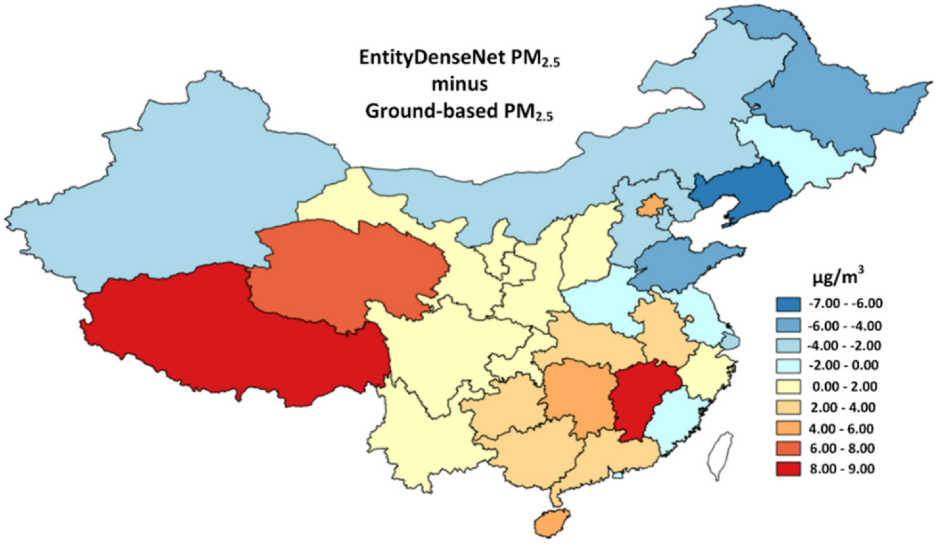

Fig. 7. The differences between provincial averaged EntityDenseNet $PM_{2.5}$ and ground-based $PM_{2.5}$ (EntityDenseNet $PM_{2.5}$ minus ground-based $PM_{2.5}$).

Table 1
Individual machine learning model performance.

Method	Training data validation		Test data validation	
	R^2	RMSE	R^2	RMSE
BPNN	0.49	31.98	0.28	38.03
XGBoost	0.64	28.33	0.57	28.69
RF	0.41	36.13	0.41	33.87
LightGBM	0.64	28.11	0.56	29.39
EntityDenseNet	**0.68**	**24.05**	**0.63**	**26.85**

EntityDenseNet. The main purpose of EntityDenseNet is to establish the nonlinear relationship between satellite spectral measurements and $PM_{2.5}$ concentration. Because the ground-based $PM_{2.5}$ mass concentration is given in an hourly scale, the 10-min resolution of the satellite spectral data and satellite geometry are averaged to the same hourly scale in the training data. The integrated training data collected from ground-based $PM_{2.5}$ and satellite data are for the period from 2016 to 2018, which contains 4,490,474 samples. The initialization scheme for EntityDenseNet is based on the Xavier initialization method (Glorot and Bengio, 2010), which confirms that the weights are not extremely small or big during the propagation of the network, and which initializes the weights from a distribution:

$$w \sim U\left[-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}\right] \quad (7)$$

where w is the weight at each layer, U is the uniform distribution in the interval $(-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}})$, and n is the size of the previous layer.

The hyper parameters in EntityDenseNet, including the value D and v in the Embedding matrix (Fig. 2), the number of neurons in the fully connected layer, and the value of the dropout rate in the dropout layer, are determined by annual 3-fold cross-validation, the details of which are shown in Fig. 3. We used the final values of the hyper parameters to train all the data (2016–2018) for the network. The collected data from 2019 (437,351 samples) was used to test the trained network system.

2.5. Extraction of spatio-temporal features from EntityDenseNet

In this study we employed the Uniform Manifold Approximation and Projection (UMAP) method (McInnes et al., 2018) to extract $PM_{2.5}$ spatio-temporal characteristics from EntityDenseNet. UMAP is a dimension reduction technique that can provide a low dimensional representation of the high dimensional feature space of an embedded categorical variable while maintaining global structures of the original feature space. The UMAP algorithm consists of two steps: creation of a graph in high dimensions followed by an optimization step to find the most similar graph in lower dimensions. To construct the initial high-dimensional graph, UMAP builds a fuzzy simplicial complex which represents a weighted graph, with edge weights representing the likelihood that two points are connected. When the fuzzy simplicial complex is constructed, UMAP projects the data into lower dimensions via a force-directed graph layout algorithm (McInnes et al., 2018).

Mapping the matrix of embedding layer from the trained EntityDenseNet down to 3D with UMAP enables us to calculate the Cosine Distance between different variables in this 3D Coordinate System. The Cosine Distance is a metric that considers the correlation of the features vector (Zou and Umugwaneza, 2008) and is defined as

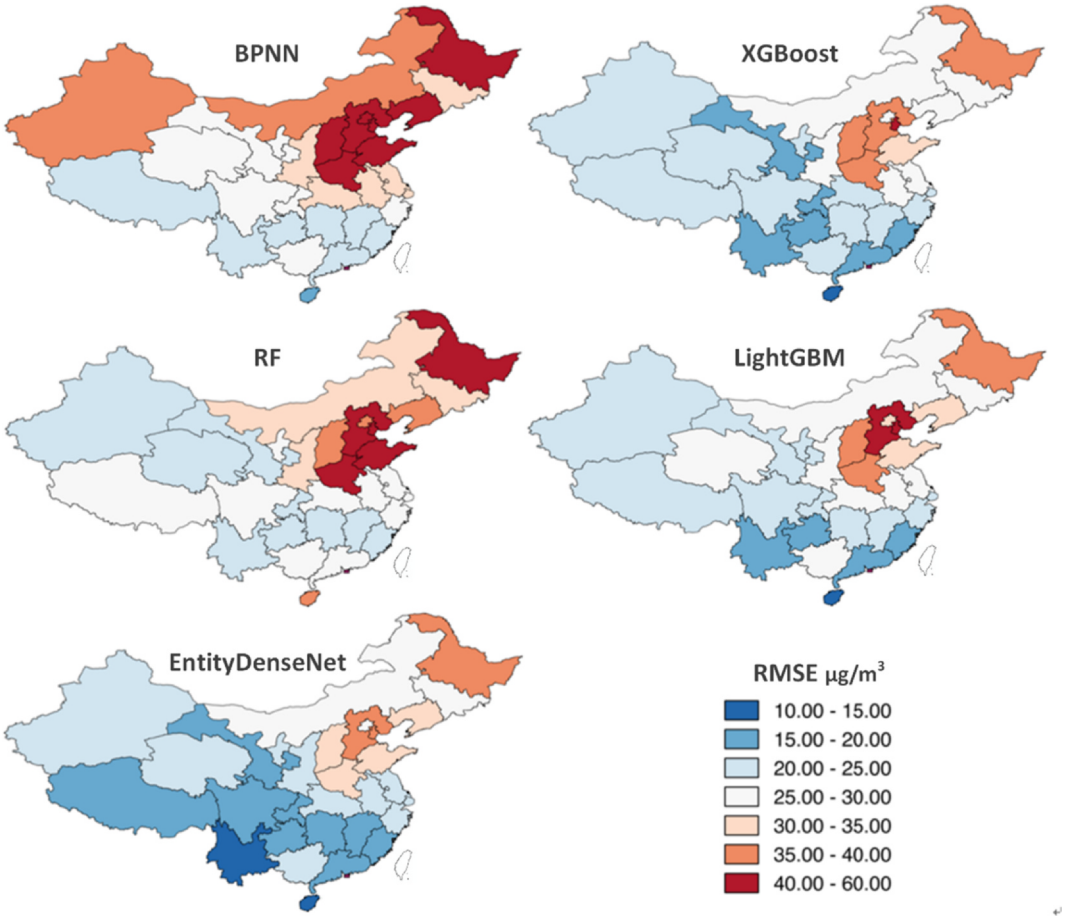

Fig. 8. Provincial RMSE values for 5 machine learning models.

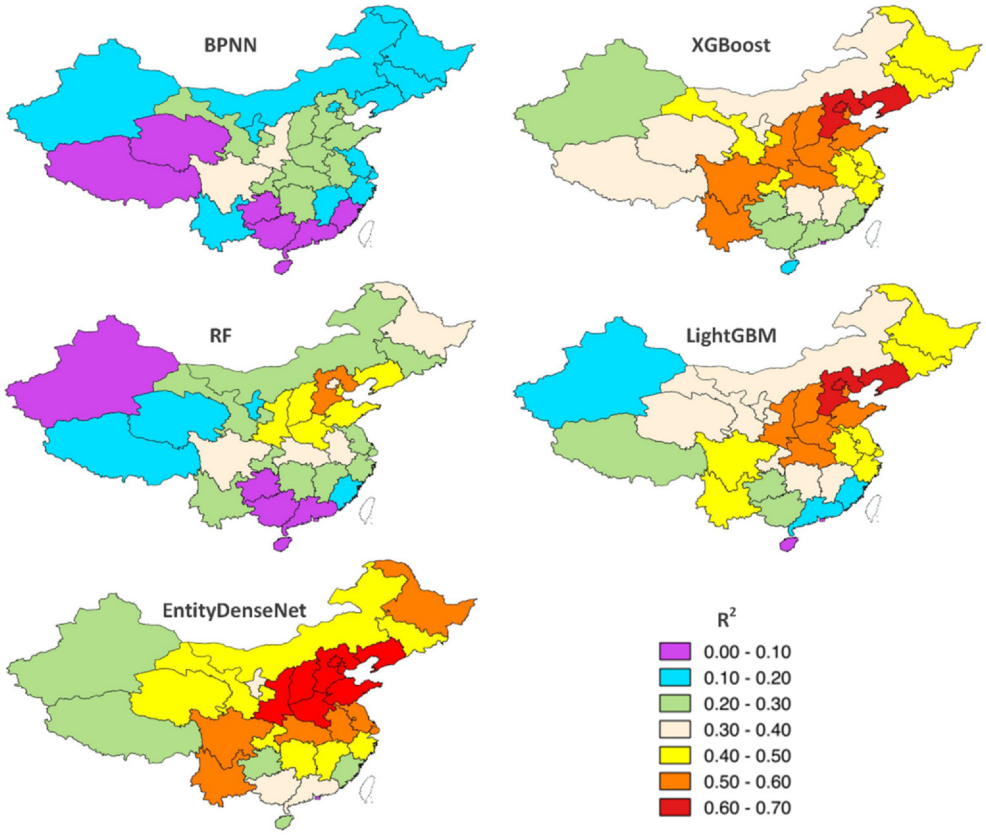

Fig. 9. Provincial R^2 values for 5 machine learning models.

$$\text{CosineDistance} = 1 - \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}} \quad (7)$$

where x_i and y_i are two vectors with length n. The smaller the Cosine Distance, the higher the correlation between two features.

2.6. Other machine learning methods

We compared four additional machine learning models to EntityDenseNet: (1) XGBoost (Chen and Guestrin, 2016), (2) BPNN (Mao et al., 2017), (3) LightGBM (Ke et al., 2017), and (4) RF (Hu et al., 2017). The hyper parameters for these methods are determined through the same training process (3-fold cross-validation) (Fig. 3), and these models used the same input data as EntityDenseNet.

3. Results

3.1. EntityDenseNet performance

Fig. 4 consist of scatterplots depicting the relationship between measured and estimated $PM_{2.5}$ for both training and test data sets. It should be noted that extremely high values were not eliminated from all the datasets to enable us to test the performance of the EntityDenseNet model in the presence of abnormal values. In our validation of the hourly test data, the linear regression between the EntityDenseNet $PM_{2.5}$ and ground-based $PM_{2.5}$ resulted in a slope of 0.56, a y-intercept of 19.45, a coefficient of determination (R^2) of 0.63, and an RMSE of 26.85 μg/m^3. As shown in Fig. 4, a small change in R^2/RMSE between training (R^2 = 0.68, RMSE = 24.05 μg/m^3) and hourly test data sets indicates a slight over-fitting in the EntityDenseNet model. For the daily and monthly averaged data validation, the RMSE values are 25.3 and 15.34 μg/m^3 while the R^2 values are 0.65 and 0.77, respectively.

The RMSE values for the EntityDenseNet estimates of $PM_{2.5}$ in each of the 1434 monitoring sites are shown in Fig. 5. Overall, 63.7% and 12.7% of the monitoring sites report RMSE values below 30 μg/m^3 and over 40 μg/m^3, respectively. The sites with RMSE values > 30 μg/m^3 are mainly concentrated in northern China, especially in the Beijing-Tianjin-Hebei (BTH) region and Shandong, whereas low RMSE values (< 20 μg/m^3) generally appear in western and southeastern China. As shown in Fig. S1, the light density corresponds to the degree of economic development, identifying the highly-developed regions of eastern China, such as the BTH region, Yangtze River Delta (YRD), and Pearl River Delta (PRD), which are the three major metropolitan areas in eastern China (Haas and Ban, 2014). Clearly, EntityDenseNet generates favorable results in PRD (RMSE mainly < 20 μg/m^3) and YRD (RMSE mainly < 30 μg/m^3), but the estimates in the BTH region are rough (RMSE > 30 μg/m^3).

Fig. 6 shows the mean $PM_{2.5}$ concentrations during the test data period from 1 January 2019 to 1 June 2019 from ground-based measurements and EntityDenseNet retrieval. The $PM_{2.5}$ spatial distribution estimates of EntityDenseNet are highly similar to those from $PM_{2.5}$ measurements. Both low value areas in western and southern China and high value areas in northern and central China are well captured. However, some overestimation still appears in central and southern China; and some areas with high values of $PM_{2.5}$ are underestimated. To better understand this bias in $PM_{2.5}$ retrieval, the differences between provincial averaged EntityDenseNet $PM_{2.5}$ and ground-based $PM_{2.5}$ (EntityDenseNet $PM_{2.5}$ minus ground-based $PM_{2.5}$) are shown in Fig. 7. The results show that overestimates of provincial averaged $PM_{2.5}$ are mainly centered in southern China, Beijing, Tibet and Qinghai, with the highest overestimation of 8.74 μg/m^3 in Tibet. Meanwhile, the underestimates of provincial averaged $PM_{2.5}$ are mainly concentrated in northern China and Xinjiang, with the largest underestimate of 6.06 μg/m^3 occurring in Liaoning. In general, the performance of EntityDenseNet $PM_{2.5}$ is better than previous studies with regard to overall

Table 2
Comparison of EntityDenseNet with other models from previous studies in China.

Model	Training data	Test data	Number of samples	Range of $PM_{2.5}$ validation (μg/m^3)	Meteorological data	AOD	Spatial resolution	Temporal resolution	R^2	RMSE (μg/m^3)	Literature
Ensemble ML model	2013–2016	2017	1,165,456	0–650	Need	Need	10 km	Daily	0.58	29	Xiao et al., 2018b
Space–time extremely randomized trees	2018	2017	177,616	0–600	Need	Need	1 km	Daily	0.65	/	Wei et al. (2020)
Space-time random forest	2015	2016	153,648	0–600	Need	Need	1 km	Daily	0.55	27.38	Wei et al. (2019)
Deep belief network	2016	CV*	4,181	0–260	Need	No Need	1 km	Daily	0.64	16.75	Shen et al. (2018)
Space-time regression modeling	2015	2014	25,601	0–680	Need	Need	10 km	Daily	0.47	37.57	He and Huang (2018)
ML + GAM	2013–2016	CV#	407,049	0–600	Need	Need	10 km	Daily	0.61	27.8	Xue et al.(2019)
EntityDenseNet	2016–2018	2019	437,351	0–1200	No Need	No Need	5 km	Hourly	0.63	26.85	This study

GAM is generalized additive model, CV# is annually iterated cross-validation, CV* is sample-based cross-validation.

bias. Lyu et al. (2017) indicated that the mean bias of estimated $PM_{2.5}$ is approximately 10 μg/m^3 in western, southeastern and northeastern China using the Community Multiscale Air Quality (CMAQ) model. However, the bias of EntityDenseNet $PM_{2.5}$ is 8.74 μg/m^3, 6.06 μg/m^3 and 4 μg/m^3 in these three regions, respectively. Xiao et al., 2018b employed an ensemble machine learning model and showed that the estimated $PM_{2.5}$ has large residuals in most stations in YRD (about −5 μg/m^3) and northern China (> 10 μg/m^3). But in this study, EntityDenseNet performs well in YRD with a bias less than 4 μg/m^3. Additionally, the bias is typically below 4 μg/m^3 in northern China.

3.2. Comparison of EntityDenseNet with other machine learning methods

Table 1 lists the performance evaluations of five methods of $PM_{2.5}$ retrieval. In both training and test data validation, only EntityDenseNet achieves an R^2 value above 0.6 and it generates the lowest RMSE values (training: 24.05 μg/m^3, test: 26.85 μg/m^3). Next to EntityDenseNet, XGBoost also performs better than the other methods in terms of RMSE (training: 28.33 μg/m^3, test: 28.69 μg/m^3). Notably, in the test data validation, the R^2 of LightGBM is similar to that of XGBoost, but its RMSE (29.39 μg/m^3) is slightly higher than that of XGBoost. From Table 1, the performance estimates of the BPNN (training: R^2 = 0.49/RMSE = 31.98 μg/m^3, test: R^2 = 0.28/RMSE 38.03 μg/m^3) and RF (training: R^2 = 0.41/RMSE = 36.13 μg/m^3, test: R^2 = 0.41/RMSE 33.87 μg/m^3) methods indicate unsatisfactory performances compared to the other methods, suggesting that these two methods lack the capacity to model the nonlinear relationship between satellite spectral measurements and ground-based $PM_{2.5}$.

The RMSE and R^2 estimates based on validation data sets from 31 provinces of China are shown Figs. 8 and 9, respectively (the detailed values in different provinces is shown in Supplementary Tables S1 and S2). In all five methods, high RMSE values (> 30 μg/m^3) are associated with the Beijing-Tianjin-Hebei (BTH) regions and northeastern China. Low RMSE values (< 15 μg/m^3) are generally observed in EntityDenseNet estimates in most provinces of mainland China. Among the different methods, RMSE levels derived from BPNN are high in 16 provinces (RMSE > 30 μg/m^3). Fig. 9 presents the distribution of local R^2 values, which can be used to examine the spatial heterogeneity and performance of each model. EntityDenseNet identifies the most number of provinces with high R^2 values (R^2 > 0.6, 6 provinces), followed by XGBoost and LightGBM (3 provinces). In contrast, BPNN identifies no province with R^2 > 0.6, which indicates that this method has the worst fit between the ground-based $PM_{2.5}$ and the estimated $PM_{2.5}$. In general, EntityDenseNet performs the best at predicting $PM_{2.5}$ among all the five models, in terms of RMSE and R^2 values over the whole of China or in 31 selected provinces. The local R^2 distributions reported by Ma et al. (2014) reveal R^2 values of 0.48–0.54 in BTH, 0.56–0.60 in YRD, > 0.58 in PRD, and 0.54–0.58 in the SC region. Lyu et al. (2017) reported R^2 values of 0.4–0.6 in BTH, 0.4–0.6 in YRD, 0.3–0.5 in PRD, and 0.3–0.5 in the SC region. In contrast to these results, the R^2 values from EntityDenseNet reflect better model performance in BTH (0.6–0.7) and the SC region (0.4–0.6). Similar levels of performance were observed in YRD (0.4–0.6) but R^2 values were low in PRD (0.3–0.4).

EntityDenseNet estimates were also compared with previous studies (Table 2). The majority of the selected studies are based on the yearly-based validation (Table 2), indicating that the test data is completely independent of the training dataset. The results suggest that although there were no input meteorological data and AOD, EntityDenseNet estimates were as good as or better than previous methods in terms of model accuracy. For example, the model proposed by Xiao et al., 2018b reported an R^2 of 0.58 for their Ensemble ML but our study improved this value to 0.63 in the EntityDenseNet estimates. The extremely high ground-based $PM_{2.5}$ values were not eliminated, and EntityDenseNet can still achieve the estimation uncertainty at 26.85 μg/m^3.

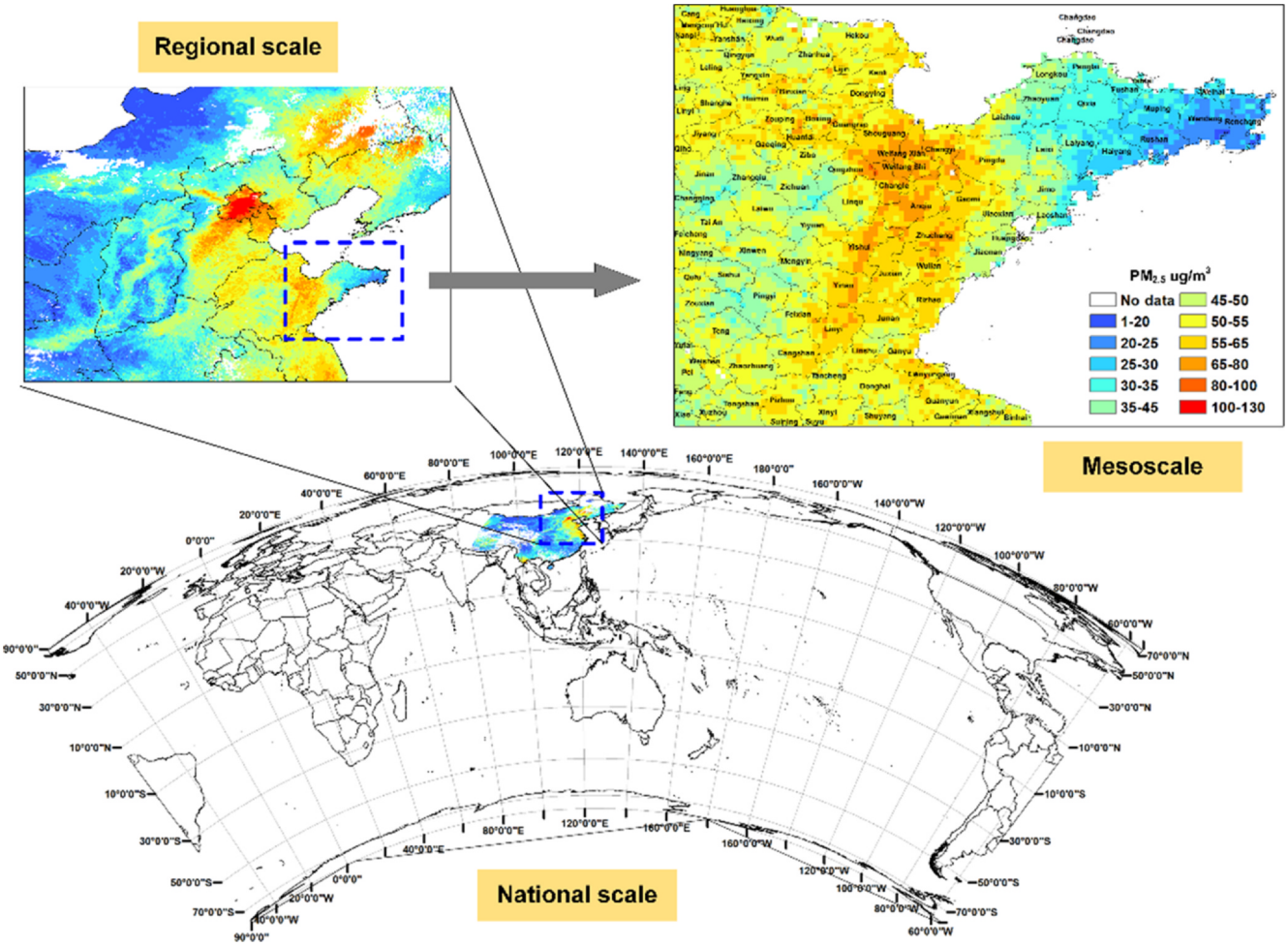

Fig. 10. Output of EntityDenseNet-based $PM_{2.5}$ concentration in different scales of view.

3.3. EntityDenseNet application

Fig. 10 presents the output of EntityDenseNet with mainland China as the area of coverage. Currently, the speed of retrieval of one full satellite image of $PM_{2.5}$ concentration over mainland China for EntityDenseNet is 5 s (CPU: Intel I3, 3.60 GHz; Memory: 16 GB; GPU: NVIDIA GeForce GTX 1060 6 GB). The results of EntityDenseNet reveal that the spatial characteristics of $PM_{2.5}$ are viewed well not only at a national scale, but also in regional and mesoscale views. In contrast to previous studies that have usually focused on retrieving the $PM_{2.5}$ concentrations at hourly (Mao et al., 2017), daily (Shen et al., 2018), monthly (Xiao et al., 2018a), seasonal (van Donkelaar et al., 2015) or annual (Lin et al., 2018) scales of temporal resolution. The temporal resolution of $PM_{2.5}$ retrieved by EntityDenseNet can be up to 10 min, which much more accurately depicts how $PM_{2.5}$ concentrations vary with time. Fig. 11 shows a case of EntityDenseNet capturing the subtle variation of $PM_{2.5}$ concentrations; the right panels depict the general hourly variation of the $PM_{2.5}$ concentrations in southern Jiangsu (circled region), where the $PM_{2.5}$ concentration dropped significantly from 75–95 μg/m^3 to 60–75 μg/m^3 on 12:00 to 13:00, April 17, 2019. The left panels of Fig. 11 show $PM_{2.5}$ concentrations every 10 min during the same period, and illustrates the $PM_{2.5}$ concentrations in southern Jiangsu dropping slightly in the first half hour (12:00–12:30) and then dramatically decreasing in the next 10 min (12:30–12:40), and later decreasing slowly again until the end (12:40–13:00). Hence, EntityDenseNet $PM_{2.5}$ with its fine temporal resolution, estimates not only the general variation of $PM_{2.5}$, but it also captures sudden changes in $PM_{2.5}$ concentration. In addition, the results from EntityDenseNet are produced without a long time lag, thus generating near real-time surface $PM_{2.5}$ concentrations at a certain spatial scale. However, it should be noted that due to the lack of ground-based measurements data at the 10 min scale, the 10-min EntityDenseNet $PM_{2.5}$ product is not well validated compared to the hourly-level product.

Fig. S2 presents three cases of the application of EntityDenseNet, including true color maps (left), real local time (13:00) (middle), and daily averaged (right) $PM_{2.5}$ distributions in mainland China on April 15–17, 2019. The true color maps show the satellite image in this period including the real cloud covering over China. In cloud-free regions, the $PM_{2.5}$ spatial distributions as well as their temporal variations are well-retrieved. Northern China also showed high concentrations of $PM_{2.5}$ during the same period. Another hotspot of high $PM_{2.5}$ concentrations is Yunnan, where $PM_{2.5}$ concentrations can reach up to 45–55 μg/m^3. Fig. S2 confirms that EntityDenseNet can retrieve $PM_{2.5}$ data for a specific local time or daily temporal resolution covering a large spatial scale.

3.4. Spatio-temporal features extraction by EntityDenseNet

The spatial characteristics of $PM_{2.5}$ between different provinces of China extracted by EntityDenseNet is displayed in Fig. 12. The province features matrix from the trained EntityDenseNet embedding layer is mapped to 3D by UMAP (Fig. 12a). The Cosine Distance between different provinces was calculated based upon Fig. 12a. In the Beijing-Tianjin-Hebei region, Tianjin had a closer Cosine Distance with Hebei (0.26) than with Beijing (0.45). This result illustrates that the $PM_{2.5}$ in Tianjin was influenced more by the $PM_{2.5}$ from Hebei than Beijing

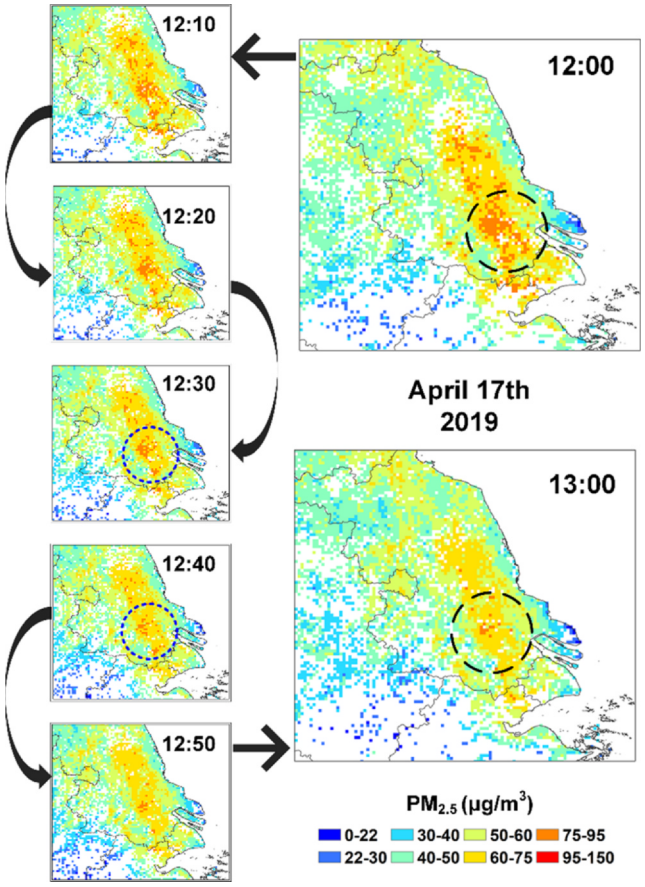

Fig. 11. Application of EntityDenseNet on April 17, 2019 from 12:00 to 13:00 (local time) over southern Jiangsu, China.

(Fig. 12a,b). This result is consistent with the WRF-Chem modeling outcome which demonstrated that $PM_{2.5}$ from Hebei had a greater contribution to Tianjin's $PM_{2.5}$ than Beijing's (Meng et al., 2020). Xue et al. (2014) also indicated that the contribution of Hebei to the $PM_{2.5}$ of Tianjin is 26%, while its contribution to Beijing is 24%. In Jiangxi, China, a closer correlation with the $PM_{2.5}$ from Hunan (Cosine Distance = 0.51) than the Hubei's $PM_{2.5}$ (Cosine Distance = 0.54) or other neighboring provinces (Cosine distance > 0.6) was identified. According to source apportionment of the $PM_{2.5}$ in Jiangxi, other provinces contributed approximately 48% of $PM_{2.5}$ annually (Xue et al., 2014). In the present study, EntityDenseNet further indicated that the Jiangxi $PM_{2.5}$ was most closely associated with Hunan. From Fig. 12, EntityDenseNet improves the understanding of the impact of $PM_{2.5}$ pollution on the provincial scale.

The month features matrix from the trained EntityDenseNet embedding layer is mapped by UMAP (Fig. 13). Although meteorological data were not used in retrieving $PM_{2.5}$, EntityDenseNet can still detect the seasonal features of $PM_{2.5}$ over China. Three groups of months with the closest Cosine Distance were observed: December, January and February (DJF), March, April and May (MAM), as well as July, August and September (JAS). DJF and MAM are winter and spring months in China and have similar meteorological conditions. Similarity level of $PM_{2.5}$ was detected by EntityDenseNet in these two seasons even without temperature and humidity information. While the group JAS are not technically deemed summer months in China, the interpretation by EntityDenseNet revealed that the $PM_{2.5}$ is more closely related in JAS, rather than in the summer months of June, July and August.

4. Discussion

Fig. 4 shows a few extreme $PM_{2.5}$ values that may be due to malfunctioning $PM_{2.5}$ monitors. In total, extremely high $PM_{2.5}$ values (> 400 μg/m^3) comprise 0.07% (N = 4,490,474) and 0.05% (N = 437,351) of the total observations of the training and test data sets, respectively. Significant underestimation of $PM_{2.5}$ may be due to local $PM_{2.5}$ sources such as coal burning, atmospheric dust, or the influence of marine salt in coastal provinces (Luo et al., 2017). Also, as the test data were gathered during winter and spring when many provinces in northern China experience snow, low-level clouds, and frequent haze days, all of which may result in $PM_{2.5}$ underestimations in northern China (Geng et al., 2015). In addition, it is possible that very low PBLH may constrain $PM_{2.5}$ to the lower atmosphere, thus shortening the path length for satellite observation and increasing the uncertainty between satellite spectral measurements and $PM_{2.5}$ (Gupta and Christopher, 2009).

The BTH region is one of the most populated and polluted regions in China due to its long-term rapid economic development and industrialization (Ma et al., 2016a, 2016b). The RMSE for EntityDenseNet $PM_{2.5}$ in this region is relatively high as shown in Fig. 5. High RMSE values in the BTH region agree with results of previous studies. For example, Zheng et al. (2015) estimated the RMSE of three major metropolitan areas in China (BTH region, YRD, and PRD), and found that the RMSE of the BTH region is ~1.3 and ~1.9 times higher than that of the YRD and PRD regions, respectively. Chen et al. (2018) calculated the provincial RMSE in China from three methods: the RF Model, Generalized Additive Model (GAM) and Non-linear Exposure-Lag-Response Model (NEM), and found that the RMSE values of both GAM and NEM are the highest in BTH region. From these results, we suggest that the accurate retrieval of $PM_{2.5}$ in the BTH region by satellite data is an issue that needs further exploration.

In the EntityDenseNet application, the only source of input data is satellite data. This feature means that it can retrieve real-time data on the spatial distribution of $PM_{2.5}$ when the satellite data is available (the JAM releases the Himawari-8 data in real-time). Compared with other machine learning methods (Table 1), EntityDenseNet provides a better solution for estimating ground-level $PM_{2.5}$ concentrations from satellite measurements data. This because with EntityDenseNet, we have introduced many features and technologies that improve the capacity to describe the nonlinear relationship between satellite data and $PM_{2.5}$. In particular: (1) the Entity Embeddings method can learn the informative relationship between categorical variables (Guo and Berkhahn, 2016); (2) a dropout layer (Srivastava et al., 2014) for each hidden layer in the EntityDenseNet has been introduced to reduce the overfitting problem; (3) the problems of saturation and vanishing gradients are overcome by using the ReLU as the activation function (Nair and Hinton, 2010); and (4) the data between the inputs in EntityDenseNet are normalized by BN technology (Ioffe and Szegedy, 2015), which fixes the mean and variance of the inputs to accelerate the training process. As shown in Table 1, when processing the same information, EntityDenseNet is more accurate and reliable than BPNN, XGBoost, LightGBM, and RF. Moreover, the EntityDenseNet can map similar values close to each other within the 3D space by UMAP (McInnes et al., 2018) and this reveals the intrinsic properties of the categorical variables. This ability greatly improves interpretability of the EntityDenseNet inversion result. EntityDenseNet can detect the $PM_{2.5}$ relationship between different provinces and months (Figs. 12 and 13). This form of deep learning approach-based spatio-temporal analysis could be extremely valuable for other earth data and scenarios (refer to Reichstein et al., 2019). However, the spatio-temporal features of $PM_{2.5}$ are complicated, as these may be influenced by other pollutants. Hence, further studies are required to comprehensively validate the interpretation of the EntityDenseNet.

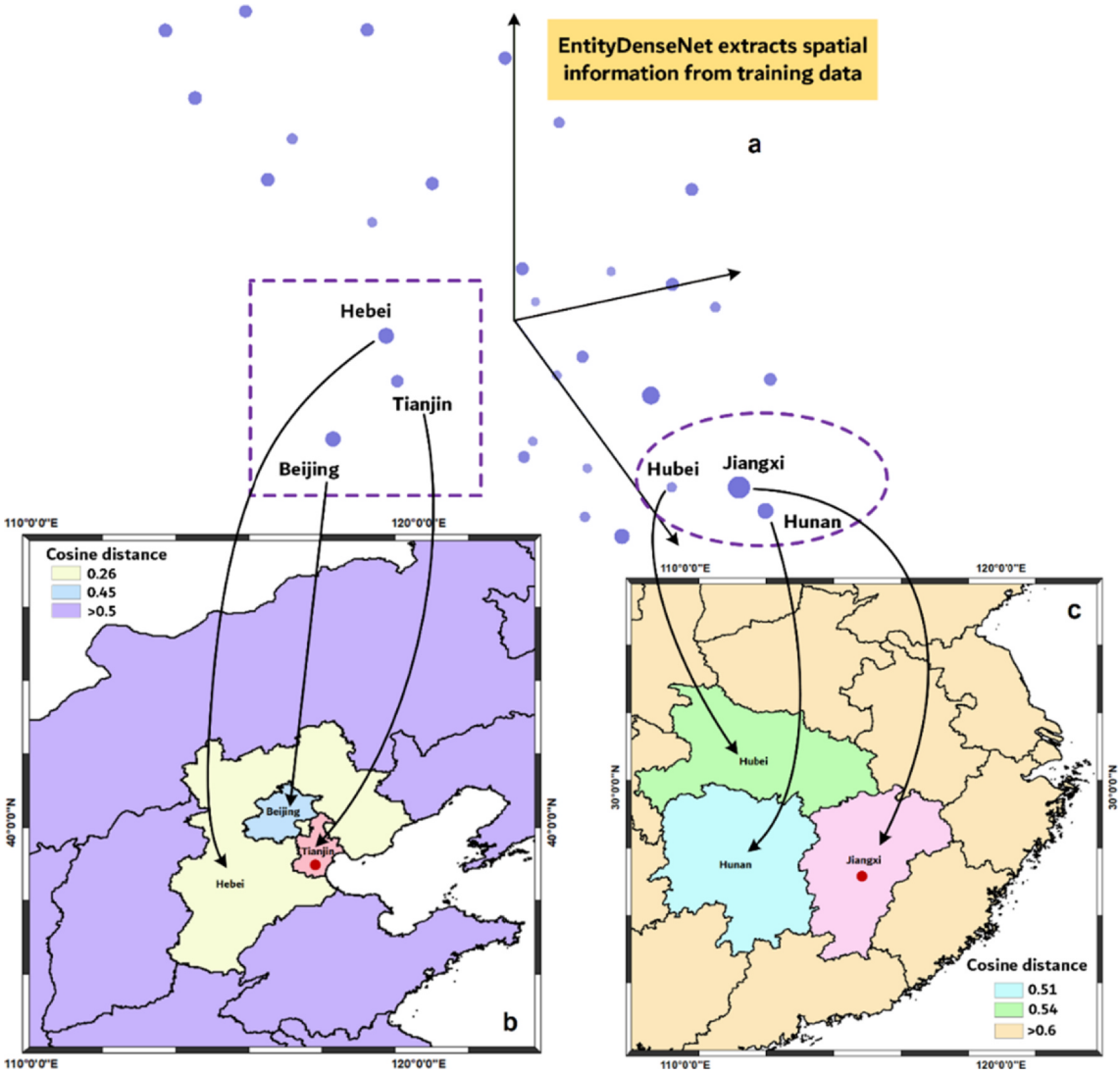

Fig. 12. Spatial analysis from EntityDenseNet. (a) The province features matrix from the trained EntityDenseNet embedding layer is mapped to 3D by UMAP. (b) The Cosine Distance to Tianjin, China. (c) The Cosine Distance to Jiangxi, China.

5. Conclusions

This study developed an interpretable deep learning model named EntityDenseNet for real-time ground-level retrieval of $PM_{2.5}$ concentrations from satellite measurement data. The $PM_{2.5}$ were retrieved from Himawari-8 observations over mainland China in 2019. As validated with ground-based measurements, the results obtained by EntityDenseNet reflected a good retrieval capability at hourly, daily and monthly time scales with RMSE values of 26.85, 25.3 and 15.34 $\mu g/m^3$, respectively. Compared to BPNN, XGBoost, LightGBM, and RF the performance of EntityDenseNet on a common set of training and test data shows the best correlation (R^2) and smallest difference (RMSE) values with the true $PM_{2.5}$ data. In addition, EntityDenseNet can outperform most models presented in previous studies for both temporal resolution and predictive accuracy. The spatial features of $PM_{2.5}$ interpreted by EntityDenseNet demonstrates that in the Beijing-Tianjin-Hebei area, the $PM_{2.5}$ in Tianjin is more subject to impacts from Hebei than Beijing, which is consistent with previous studies (Meng et al., 2020; Xue et al., 2014). This relationship was further explored using the deep learning approach in the present study. Without meteorological data, EntityDenseNet can still detect the seasonal features of $PM_{2.5}$ over China, which indicated that the $PM_{2.5}$ appears increasingly similar in three month groups over mainland China: (1) December, January and February, (2) March, April and May, (3) July, August and September.

This study revealed that a new, deep learning model achieves higher accuracy when retrieving real-time ground-level $PM_{2.5}$ concentrations directly from satellite spectral data. Further, it has the interpretable ability to extract spatio-temporal features of $PM_{2.5}$. Owing to the high temporal resolution of Himawari-8, this new, deep learning model can estimate large-scale $PM_{2.5}$ concentrations at hourly intervals. This will improve our understanding of how $PM_{2.5}$ concentrations vary in fine temporal resolution. We have created a EntityDenseNet Cloud Platform (http://49.233.1.40:8888/), it is free to access and researchers can use it for their own data modeling. The guide of this Cloud Platform is shown in the Supplementary material.

CRediT authorship contribution statement

Xing Yan: Conceptualization, Methodology, Software, Writing - original draft, Supervision. **Zhou Zang:** Investigation, Validation, Writing - original draft. **Nana Luo:** Methodology, Writing - original draft. **Yize Jiang:** Methodology, Software. **Zhanqing Li:** Funding acquisition, Project administration, Writing - review & editing.

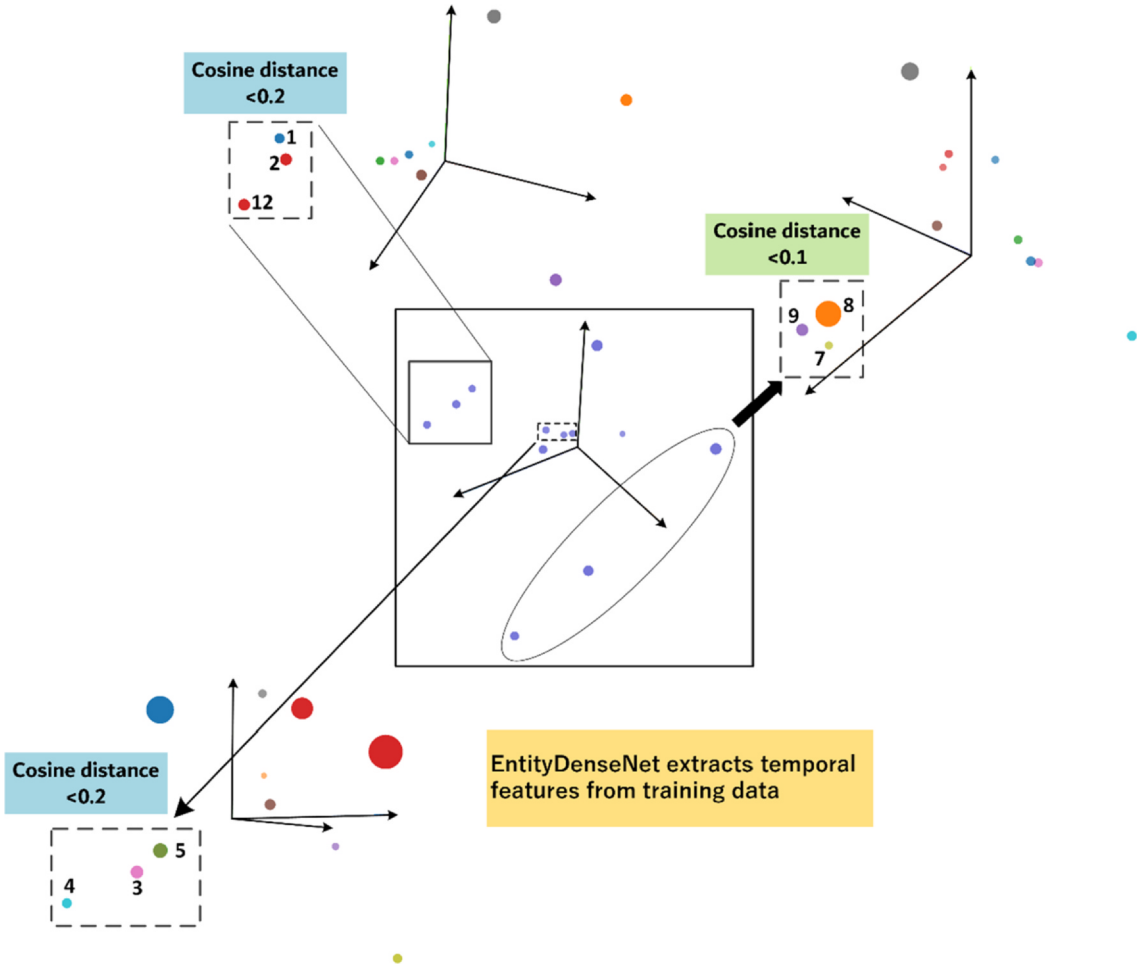

Fig. 13. Temporal analysis from EntityDenseNet. The month features matrix from the trained EntityDenseNet embedding layer is mapped to 3D by UMAP.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the National Key Research and Development Plan of China (2017YFC1501702), the National Natural Science Foundation of China (91544217, 41801329, 91837204), and the Fundamental Research Funds for the Central Universities. The authors gratefully acknowledge the Japan Meteorological Agency (JMA) for providing Himawari-8 data.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.106060.

References

Amanollahi, J., Tzanis, C., Abdullah, A., Ramli, M., Pirasteh, S., 2013. Development of the models to estimate particulate matter from thermal infrared band of Landsat Enhanced Thematic Mapper. Int. J. Environ. Sci. Technol. 10, 1245–1254.

Apte, J.S., Marshall, J.D., Cohen, A.J., Brauer, M., 2015. Addressing Global Mortality from Ambient PM2.5. Environ. Sci. Technol. 49 (13), 8057–8066.

Ben Khalifa, A., Gazzah, S., EssoukriBenAmara, N., 2013. Adaptive score normalization: a novel approach for multimodal biometric systems. Int. J. Comput. Inf. Sci. Eng 7 (3), 18–26.

Chen, G., Li, S., Knibbs, L., Ham, N.A.S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M., Guo, Y., 2018. A machine learning method to estimate PM2.5 concentrations across China with remote sensing, meteorological and land use information. Sci. Total Environ. 636, 52–60.

Chen, T., Guestrin, C., 2016, August. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM, pp. pp. 785–794.

Chren, W.A., 1998. One-hot residue coding for low delay-power product CMOS design. IEEE T CIRCUITS-II 45 (3), 303–313.

Geng, G., Zhang, Q., Martin, R., Donkelaar, A., Huo, H., Che, H., Lin, J., He, K., 2015. Estimating long-term PM2.5 concentrations in China using satellite-based aerosol optical depth and a chemical transport model. Remote Sens. Environ. 166, 262–270.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. J. Mach. Learn. Res. – Proc. Track 9, 249–256.

Guo, C., Berkhahn, F., 2016. Entity embeddings of categorical variables. arXiv preprint arXiv:1604.06737.

Guo, Y., Feng, N., Christopher, S.A., Kang, P., Zhan, F.B., Hong, S., 2014. Satellite remote sensing of fine particulate matter (PM2.5) air quality over Beijing using MODIS. Int. J. Remote Sens. 35 (17), 6522–6544.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: a highly efficient gradient boosting decision tree. Adv. Neur. Inf. Process.Sys. 30, 3146–3154.

Gupta, P., Christopher, S.A., 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. J. Geophys. Res. Atmos. 114 (D14).

Gupta, P., Christopher, S.A., Wang, J., Gehrig, R., Lee, Y., Kumar, N., 2006. Satellite remote sensing of particulate matter and air quality assessment over global cities. Atmos. Environ. 40 (30), 5880–5892.

Haas, J., Ban, Y., 2014. Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. Int. J. Appl. Earth Obs. Geoinf. 30, 42–55.

Han, Y., Wu, Y., Wang, T., Zhuang, B., Li, S., Zhao, K., 2015. Impacts of elevated-aerosol-layer and aerosol type on the correlation of AOD and particulate matter with ground-based and satellite measurements in Nanjing, southeast China. Sci. Total Environ. 532, 195–207.

He, Q., Huang, B., 2018. Satellite-based mapping of daily high-resolution ground PM2. 5 in China via space-time regression modeling. Remote Sens. Environ. 206, 72–83.

Hoff, R.M., Christopher, S.A., 2009. Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land? J. Air Waste Manage. Assoc. 59 (6), 645–675.
Hu, H., Hu, Z., Zhong, K., Xu, J., Zhang, F., Zhao, Y., Wu, P., 2019. Satellite-based high-resolution mapping of ground-level PM2.5 concentrations over East China using a spatiotemporal regression kriging model. Sci. Total Environ. 672, 479–490.
Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L., Strickland, M., Liu, Y., 2017. Estimating PM2.5 Concentrations in the Conterminous United States Using the Random Forest Approach. Environ. Sci. Technol. 51 (12), 6936–6944.
Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167.
Ishida, H., Nakajima, T.Y., 2009. Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. J. Geophys. Res. Atmos. 114 (D7).
Koelemeijer, R., Homan, C.D., Matthijsen, J., 2006. Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. Atmos. Environ. 40, 5304–5315.
Kokhanovsky, A.A., Prikhach, A.S., Katsev, I.L., Zege, E.P., 2009. Determination of particulate matter vertical columns using satellite observations. Atmos. Meas. Tech. 2 (2), 327–335.
Levy, R.C., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., Hsu, N.C., 2013. The Collection 6 MODIS aerosol products over land and ocean. Atmos. Meas. Tech. 6 (11), 2989–3034.
Liang, C., Zang, Z., Li, Z., Yan, X., 2020. An Improved Global Land Anthropogenic Aerosol Product Based on Satellite Retrievals From 2008 to 2016. IEEE Geosci. Remote Sens. Lett. https://doi.org/10.1109/LGRS.2020.2991730.
Lin, C.Q., Liu, G., Lau, A.K.H., Li, Y., Li, C.C., Fung, J.C.H., et al., 2018. High-resolution satellite remote sensing of provincial PM2.5 trends in China from 2001 to 2015. Atmos. Environ. 180, 110–116.
Liu, H., Hussain, F., Tan, C.L., Dash, M., 2002. Discretization: An Enabling Technique. Data Min. Knowl. Discov. 6, 393–423.
Liu, Y., Cao, G., Zhao, N., Mulligan, K., Ye, X., 2017. Improve ground-level PM2.5 concentration mapping using a random forests-based geostatistical approach. Environ. Pollut. 235, 272–282.
Liu, Y., Paciorek, C., Koutrakis, P., 2009. Estimating Regional Spatial and Temporal Variability of PM2.5 Concentrations Using Satellite Data, Meteorology, and Land Use Information. Environ. Health Perspect. 117, 886–892.
Luo, J., Du, P., Samat, A., Xia, J., Che, M., Xue, Z., 2017. Spatiotemporal Pattern of PM2.5 Concentrations in Mainland China and Analysis of Its Influencing Factors using Geographically Weighted Regression. Sci. Rep. 7, 40607.
Lyu, B., Zhang, Y., Hu, Y., 2017. Improving PM2.5 Air Quality Model Forecasts in China Using a Bias-Correction Framework. Atmosphere. 8, 147.
Ma, X., Wang, J., Yu, F., Jia, H., Hu, Y., 2016a. Can MODIS AOD be employed to derive PM2.5 in Beijing-Tianjin-Hebei over China? Atmos. Res. 181, 250–256.
Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating Ground-Level PM2.5 in China Using Satellite Remote Sensing. Environ. Sci. Technol. 48, 7436–7444.
Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., et al., 2016. Satellite-based spatiotemporal trends in PM2. 5 concentrations: China, 2004–2013. Environ. Health Perspect., 124(2), 184–192.
Mao, X., Shen, T., Feng, X., 2017. Prediction of hourly ground-level PM2.5 concentrations 3 days in advance using neural networks with satellite data in eastern China. Atmos. Pollut. Res. 8, 1005–1015.
McInnes, L., Healy, J., Melville, J., 2018. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
Meng, L., Cai, Z., Li, Y., et al., 2020. Spatial and temporal distributions and source simulation during heavy pollution of PM2.5 in Tianjin. Res. Environ. Sci. 1, 9–17.
Nair, V., Hinton, G. E., 2010. Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807–814.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566 (7743), 195–204.
Sayer, A., Munchak, L., Hsu, N., Levy, R., Bettenhausen, C., Jeong, M.J., 2014. MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and 'merged' datasets, and usage recommendations. J. Geophys. Res. Atmos. 119.
Shen, H., Li, T., Yuan, Q., Zhang, L., 2018. Estimating Regional Ground-Level PM2.5 Directly From Satellite Top-Of-Atmosphere Reflectance Using Deep Belief Networks. J. Geophys. Res. Atmos. 123 (24), 13875–13886.
Sorek-Hamer, M., Kloog, I., Koutrakis, P., Strawa, A., Chatfield, R., Cohen, A., Ridgway, W., Broday, D., 2015. Assessment of PM2.5 concentrations over bright surfaces using MODIS satellite observations. Remote Sens. Environ. 163, 180–185.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958.
van Donkelaar, A., Martin, R.V., Park, R.J., 2006. Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing. J. Geophys. Res. Atmos. 111 (D21).
van Donkelaar, A., Martin, R.V., Spurr, R.J.D., Burnett, R.T., 2015. High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America. Environ. Sci. Technol. 49 (17), 10482–10491.
Wei, J., et al., 2020. Improved 1 km resolution pm2.5 estimates across china using enhanced space-time extremely randomized trees. Atmosph. Chem. Phys. 20 (6), 3273–3289.
Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019. Estimating 1-km-resolution PM2. 5 concentrations across China using the space-time random forest approach. Remote Sens. Environ. 231, 111221.
Xiao, Q., Chang, H.H., Geng, G., Liu, Y., 2018b. An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. Environ. Sci. Technol. 52 (22), 13260–13269.
Xiao, L., Lang, Y., Christakos, G., 2018a. High-resolution spatiotemporal mapping of PM2.5 concentrations at Mainland China using a combined BME-GWR technique. Atmos. Environ. 173, 295–305.
Xu, Y., Ho, H.C., Wong, M.S., Deng, C., Shi, Y., Chan, T.C., Knudby, A., 2018. Evaluation of machine learning techniques with multiple remote sensing datasets in estimating monthly concentrations of ground-level PM2.5. Environ. Pollut. 242, 1417–1426.
Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T., Zhang, Q., 2019. Spatiotemporal continuous estimates of PM2. 5 concentrations in China, 2000–2016: A machine learning method with inputs from satellites, chemical transport model, and ground observations. Environ. Int. 123, 345–357.
Xue, W., Fu, F., Wang, J., Tang, G., Lei, Y., Yang, J., Wang, Y., 2014. Numerical study on the characteristics of regional transport of PM2.5 in China. China Environ. Sci. 6, 1361–1368.
Yan, X., Li, Z., Luo, N., et al., 2019. An improved algorithm for retrieving the fine-mode fraction of aerosol optical thickness. Part 2: Application and validation in Asia. Remote Sens. Environ. 222, 90–103.
Yan, X., Liang, C., Jiang, Y., Luo, N., Zang, Z., Li, Z., 2020. A Deep Learning Approach to Improve the Retrieval of Temperature and Humidity Profiles From a Ground-Based Microwave Radiometer. IEEE Trans. Geosci. Remote Sens. https://doi.org/10.1109/TGRS.2020.2987896.
Yan, X., Shi, W., Li, Z., Li, Z., Luo, N., Zhao, W., Wang, H., Yu, X., 2017. Satellite-based PM2.5 estimation using fine-mode aerosol optical thickness over China. Atmos. Environ. 170, 290–302.
Yan, X., Shi, W., Zhao, W., Luo, N., 2014. Impact of aerosols and atmospheric particles on plant leaf proteins. Atmos. Environ. 88, 115–122.
Yan, X., Shi, W., Zhao, W., Luo, N., 2015. Mapping dustfall distribution in urban areas using remote sensing and ground spectral data. Sci. Total Environ. 506, 604–612.
Yang, X., Jiang, L., Zhao, W., Xiong, Q., Zhao, W., Yan, X., 2018. Comparison of Ground-Based PM2.5 and PM10 Concentrations in China, India, and the U.S. Int J Environ Res Public Health. 15 (7), 1382.
You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., Wang, W., 2016. National-Scale Estimates of Ground-Level PM2.5 Concentration in China Using Geographically Weighted Regression Based on 3 km Resolution MODIS AOD. Remote Sens. 8, 184.
Zhang, Y., Li, Z., 2013. Estimation of PM2.5 from fine-mode aerosol optical depth. J. Remote Sens. 17, 929–943.
Zhang, Y., Li, Z., 2015. Remote sensing of atmospheric fine particulate matter (PM2.5) mass concentration near the ground from satellite observation. Remote Sens. Environ. 160, 252–262.
Zheng, Y., Zhang, Q., Liu, Y., Geng, G., He, K., 2015. Estimating ground-level PM 2.5 concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. Atmos. Environ. 124.
Zou, B.J., Umugwaneza, M.P., 2008. Shape-based trademark retrieval using cosine distance method. In: 2008 Eighth International Conference on Intelligent Systems Design and Applications, vol. 2. IEEE, pp. 498–504.