Science of the Total Environment 724 (2020) 138178

Contents lists available at ScienceDirect

Science of the Total Environment

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

A deep learning and image-based model for air quality estimation

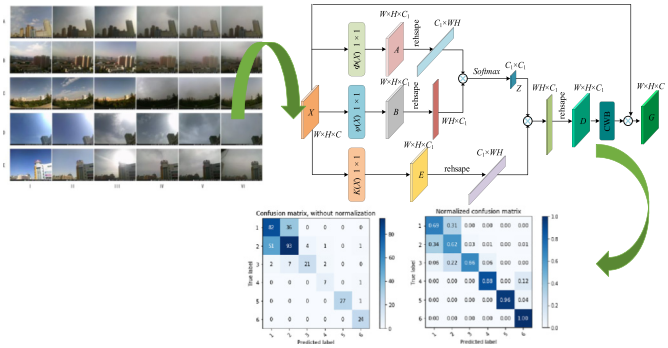
Qiang Zhang *, Fengchen Fu, Ran Tian

College of Computer Science and Engineering, Northwest Normal University, Lanzhou, Gansu Province, 730070, China

HIGHLIGHTS

- A method based on residual neural network was proposed to detect air quality from images collected by mobile devices.
- Designed the self-supervision (SCA) module, improved the air quality level detection model, and improved the recognition accuracy.
- Our method outperforms other machine learning based methods.
- Set up a data set an image data set containing 5 scenes, and mark the real-time monitoring data of the corresponding air quality monitoring station as an image

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 28 January 2020
Received in revised form 22 March 2020
Accepted 23 March 2020
Available online 24 March 2020

Editor: Jianmin Chen

Keywords:
Convolutional neural network
Public satisfaction
Participatory sensing
Residual network

ABSTRACT

The serious threat of air pollution to human health makes air quality a focus of public attention, and effective, timely air quality monitoring is critical to pollution control and human health. This paper proposes a deep learning and image-based model for air quality estimation. The model extracts feature information from scene images captured by camera equipment and then classifies them to estimate air quality levels. A self-supervision module (SCA) is added to the model and the global context information of the feature map is used to reconstruct the features by using the interdependence between the channel maps to enhance the interdependent channel maps and improve the ability of feature representation. In addition, a high-quality outdoor air quality data set (NWNU-AQI) was compiled to facilitate the training and evaluation of the model's performance. This paper compares and analyzes AQC-Net, Support Vector Machine (SVM), and Deep Residual Network (ResNet) on NWNU-AQI. The experimental results show that AQC-Net yields more accurate results for air quality classification than other methods.

1. Introduction

Long-term exposure to air pollution causes serious harm to human health. Excessive concentration of fine particles in the air causes worsening of human respiratory diseases and exacerbating respiratory infections, chronic pharyngitis, chronic bronchitis, bronchial asthma, emphysema and other respiratory diseases (Ameer et al., 2019; Portney and Mullahy, 1990), as well as affecting children's lung development (Chipps, 2015). Particulate matter also increases the risk of sudden myocardial infarction, causing ischemic heart disease, cerebrovascular disease, chronic obstructive pulmonary disease, and lung cancer (Burnett et al., 2014). Particulate matter enters the body through the respiratory tract, and can even contribute to brain damage such as brain atrophy and memory decline, increasing the risk of Alzheimer's disease (Younan et al., 2019; Zhan et al., 2018). Pollutants entering the

* Corresponding author.
E-mail address: zhangq@nwnu.edu.cn (Q. Zhang).

https://doi.org/10.1016/j.scitotenv.2020.138178
0048-9697/

human body are transformed into compounds that are more carcinogenic or mutagenic, leading to an increase in the incidence and mortality of cancer, and the number of deaths from air pollution in some areas even exceeds traffic accidents (Raaschou-Nielsen et al., 2013; Li et al., 2015a; Chipps, 2015; Dockery et al., 1993).

Air pollution not only harms the human body but also affects human life in other ways. Severe air pollution creates haze, which reduces atmosphere visibility, increases the probability of traffic accidents, and even causes flight delays, affecting people's normal travel and productivity.

The public's real-time access to air quality information (Zheng et al., 2013) and, in response, taking timely protection and prevention measures are crucial to prevent health threats caused by air pollution exposure. Therefore, air quality monitoring is essential.

2. Related work

2.1. Traditional method

Many methods of monitoring air quality have been developed over the past 40 years. These traditional methods can be categorized into four groups (De-zhong, 2013;CUI et al., 2012): filter-based weighting by manual sampling (Hauck et al., 2004); cone element oscillating microbalance methods for automatic measurement (Patashnick and Rupprecht, 1991); optical-based particle extinction and scattering methods (Knollenberg, 1970); and β radiation attenuation test (Nader, 1975). All these methods, however, have limitations. As they sample particulate matter in the air during the sampling period and then calculate the weight change a sampling filter is needed on-site for measurement. The direct measurement method mainly depends on the filter screen collection. The filter screen can only work for a period of time and needs to be replaced in time. The system needs regular maintenance and the cost is high. Manual measurement has high accuracy but its operation is complicated. Automatic measurement is more convenient but requires frequent calibration.

At present, countries around the world have set up many air quality monitoring stations in various cities to obtain pollutant concentrations through various precision detection instruments based on the above principles. These stations, calculate the air quality index (AQI) and release it to the public. The layout of the monitoring station takes into account both the population distribution and the size of the built-up area, as well as the representativeness, continuity, safety and operability of the station locations. As a result, a monitor station can only capture the air quality of within 1–3 km (Liu et al., 2010). Air quality in a city, however, varies greatly across regions, and changes non-linearly (Adams and Kanaroglou, 2016; Zheng et al., 2013), such that it cannot be adequately captured by a limited number of monitoring stations (Yong-hong, 2011). Another limitation of this method is that sensors have a short life and require regular maintenance. It is too costly to achieve dense deployment and cover every corner of a city. In short, people living and working far from monitoring stations cannot obtain their accurate and real-time air quality reports.

To more accurately represent of city-wide air quality, some researchers have proposed several methods, such as spatial averaging and recent Neighbor law (Schwartz, 1989; Chestnut et al., 1991), to utilize the limited data from monitoring stations based on spatial interpolation. These methods are mainly based on the assumption of spatial continuity in the distribution of air pollution particles and solve the problem of data sparsity by interpolating the monitoring data. However, these methods have two main shortcomings. The first is that estimates obtained by different methods are widely different. Second, for raw data with relatively sparse spatial distribution, the difference results are not ideal. Several researchers have attempted to further optimize the air quality detection network (Li et al., 2012; Rui and Ming-Shun, 2015). Mei et al. (2014) proposed a method based on mobile information to monitor air quality complementary to monitoring stations.

Researchers are increasingly focusing on crowdsourcing computing, including the use of auxiliary sensors (such as remote sensing, on-board sensors). Murty et al. (2008) proposed a wireless technology air detection network platform called CitySense to monitor air pollutants air and related indicators. Yu et al. (2012) proposed a VSN-based monitoring method to achieve data sampling through compressed sensing technology. Li et al. (2014) monitored gas pollutants and particulate matter through portable sensors and smartphones. However, current portable sensors cannot achieve the accuracy of monitoring stations. In addition, $PM_{2.5}$ monitoring equipment requires more than 1 h of data collection, and it is also necessary to avoid errors caused by movement as much as possible.

There are also studies on air quality monitoring based on satellite remote sensing. Gupta et al. (2006) estimated the global $PM_{2.5}$ quality from the aerosol optical thickness extracted from satellite imaging data. Padayachi (2016) used improved satellite aerosol remote sensing as an alternative to estimate ground PM_{10} mass concentration. Chung (1986) used satellites to detect air pollution and its transmission.

Other methods have also been proposed, such as those based on black smoke measurement (Muir and Laxen, 1995), atmospheric aerosol optical information (Smith and Atkinson, 2001), spectral technology monitoring (Hodgeson et al., 1973), data-driven spatial and temporal deep learning (Li et al., 2016), and social network (Chen et al., 2014). There is also the method that constructs a classifier based on data from monitoring stations and meteorological data, taxi trajectories, road networks, etc. through artificial neural networks (ANN) (Zheng et al., 2013).

2.2. Image-based methods

With the rapid development of smartphones and video surveillance equipment and the widespread application of artificial intelligence (Alson and Misagal, 2016; Yu et al., 2016), image quality has continued to improve, and acquisition and collection have become simpler and more convenient, making it possible to use artificial intelligence methods such as image processing and machine learning to detect the air quality in images. The public can easily take pictures of their surroundings on their mobile phones, and use established air quality image recognition models to analyze and obtain air quality information, which can inform the public to take timely air pollution protection measures. Especially, with the continuous development and application of new methods of deep learning in the field of image recognition, the recognition of air quality levels based on scene image analysis is receiving increasing attention. The use of images to detect air quality can greatly reduce the reliance on professional hardware and equipment, as well as the labor and material resources required for equipment maintenance, making it more convenient and efficient. It can also improve the spatial granularity of air quality monitoring. Image-based air quality evaluation can be roughly divided into two categories: image features based methods and deep learning based methods.

2.2.1. Image feature-based methods

These methods focus on the relationships between the characteristics of the image and the PM index. Liu et al. (2011) evaluated the correlation between air quality and image quality. Fattal and others (Fattal, 2008; Tan, 2008) proposed that the research on image defogging be extended to image-based air quality evaluation. He et al. (2010) proposed an atmospheric scattering model to extract transmittance features and transmission matrices in a dark channel Mao et al. (2014) haze factors by analyzing the atmospheric scattering model and detecting fog image through the color channel Li et al. (2015b) first used the dark channel prior to estimating the transmission matrix, examined the depth map based on the deep convolution neural field (Liu et al., 2015a), and estimated haze in photos by combining the transmission matrix and the depth map to extract the haze image. The two image features of depth map and transmission matrix were used to statistically

estimate haze levels. Liu et al. (2015b) proposed a method that obtained i) an extinction coefficient for an image taken by a smartphone through a physical model and sensors and ii) a transmittance through an atmospheric scattering model to comprehensively estimate $PM_{2.5}$ concentrations., Liu et al. (2016) proposed a method based on six factors (i.e., transmittance, sky smoothness, sky color, image contrast, image entropy, and two non-image features of the sun's vertex angle and humidity); used principal component analysis (PCA) and order Backward Feature Selection (SBFS) to optimize the feature set; and established a regression model to predict $PM_{2.5}$ levels in combination with photo shooting time, geographic location, and weather conditions. Wang et al. (2014) studied air quality by analyzing the relationship between PM2.5 and observed image degradation. Zhang et al. (2015) used decision tree and multi-core learning to estimate air quality.

2.2.2. Deep learning-based methods

Deep learning-based methods have made great breakthroughs in image feature learning, and also advances in solving typical computer vision problems (Krizhevsky et al., 2012). In 2016, Zhang et al. (2016a) used a CNN-based method to estimate air pollution from photos. The constructed CNN was used to classify images according to the $PM_{2.5}$ index of the image. The CNN had 9 convolutional layers, 2 pooling layers, and 2 dropout layers, and used an improved rectified linear unit as the activation function to solve the gradient disappearance problem, and used a negative log-log ordinal classifier with graph Softmax classifier to solve the air pollution problem. The method proposed by Avijoy in (Chakma et al., 2017) applied the VGG-16CNN model to the analysis of $PM_{2.5}$ levels in the pre-box, using a transfer learning strategy (CNN fine-tuning and random forest based on CNN features) and based on the PM2.5 concentration levels of the image. Bo et al. (2018) combined image and weather information to estimate PM2.5 index of outdoor images using deep learning and support vector regression (SVR) technology Ma et al. (2018). Proposed an air quality classification method using a single image of a hybrid CNN. Their work was based on the dark channel prior theory in the defogging method proposed by He et al. (2010). Rijal et al. (2018) used a feedforward neural network combined with a convolutional neural network to estimate $PM_{2.5}$. Chen et al. (2016) input boundary pixel ratio and highlight pixel ratio features to BP neural network training to build an AQI estimation model. Pan et al. (2017) extracted the transmittance map of images as input to develop an AQI estimation model.

Air quality research based on image detection is realized by combining image processing methods and machine learning methods, but both have certain shortcomings. For example, the features used in the $PM_{2.5}$and PM_{10} concentration detection methods based on image visual characteristics can be affected by the color characteristics of the sky. The sensitivity is too high and it is greatly affected by the weather. The method of detecting $PM_{2.5}$ and PM_{10} concentrations based on physical characteristics can yield relatively results, but it is only suitable for images collected in dry air, which is affected by meteorological conditions.

To address these shortcomings, this paper proposes a new image- and deep learning-based method that i) is not limited by locations and is suitable for images collected at any locations, and ii) is timely, with the ability to use images at the time they are taken. This method transforms camera-enabled smart mobile devices into air quality sensors to accurately assess air quality using image analysis and deep learning technologies.

The method proposed in this article is different from existing methods. It detects the surrounding air pollution using images collected by mobile devices in daily life. It can make up for the shortcomings of existing air quality detection technologies and achieve fine-grained, low-cost monitoring of air quality. Compared with those methods based on image features, it uses the ResNet-based deep convolutional neural network model AQC-Net. The residual network is designed for image recognition and is not completely suitable for air quality image evaluation. This paper aims at addressing the problems of image object recognition and air quality recognition tasks with different focuses, and improving the residual network (Wen-wen, 2019). By integrating relevant features between all channel maps, the inter-dependent channel maps are enhanced to improve the representation capabilities of the feature maps. Based on a clear end-to-end architecture that can automatically extract low-level and high-level image features, the method in this paper avoids complex feature extraction and feature optimization steps. Also, existing methods based on convolutional neural networks focus almost exclusively on $PM_{2.5}$, but $PM_{2.5}$ is only part of air pollution and cannot fully reflect air quality. The proposed method can directly estimate the AQI, which is more comprehensive and better reflect air quality. This paper explores the relationship between air quality and image characteristics based on the analysis of air quality of multiple fixed scene images, establishes a prediction model, and then estimates air quality at any place. The public can easily obtain images through portable terminals such as smart phones, and immediately predict the AQI using this method.

The main contributions of this article are summarized as follows.

An image data set containing 5 scenes was collected and the images were labeled with real-time monitoring data from corresponding air quality monitoring stations. The dataset was collected from January 2018 to December 2019, with photos taken in various weather and seasons and during various time periods of a day.

This paper proposes a deep convolutional neural network model AQC-Net based on ResNet. By adding the self-supervision module SCA, the accuracy of ResNet for air quality image classification is significant improved. Experimental results show that the method is feasible.

3. Methods and analysis

3.1. Spatial and context attention block

The structure of our model is shown in Fig. 1. The training set of the model is$\{(x_i,y_i)\}_{i=1}^{N}$, $x_i \in \mathbb{R}^{H\times W\times C}$, $y_i \in N$. We define the evaluation air

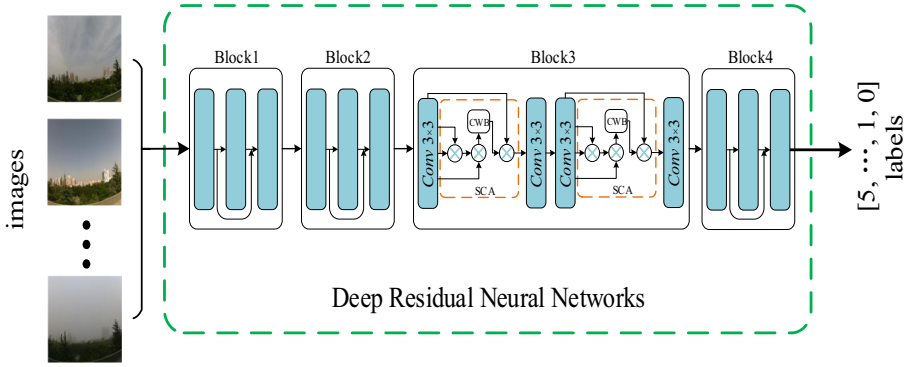


Fig. 1. Air quality classification model.

quality image set as$\{x_i\}_{i=1}^N$, and set of air quality rating labels as$\{y_i\}_{i=1}^N$. When inputting the image x_i, it is desired to obtain the air quality level y_i corresponding to x_i and the mapping relationship $y_i = F(x_i)$.

A self-supervision module is designed, called Spatial and Context Attention block (SCA), and this module is added to the original ResNet18 (He et al., 2016) network structure to form AQC-Net for feature extraction of air quality images. The residual module is more sensitive to changes in weight. The self-monitoring module can continuously adjust the importance of feature information, so that the model can better approach the global optimal solution. Keeping this original structure unchanged, a scene self-supervision module was added to the third module. The third block contains two residual structures. In the SCA module added in each residual structure, the feature map of the input SCA module in the third module has a resolution of 1/16 of the original input image, which can greatly reduce the amount of calculation during matrix multiplication. As shown in Fig. 1, the feature map generates different contextual scene feature information after three branches. The first branch is to obtain the correlation between each pixel in the air quality image, and then matrix-multiply the output of the first branch by the output of the second branch to obtain the similarity between different channel maps. Finally, the feature map generated by the third branch is matrix-multiplied by the above result to redistribute the relevant feature information to the original feature map, so as to obtain the correlation between the information of the entire feature map. Finally, the feature map is aggregated using global average pooling (Lin et al., 2013) and multiplied by the input feature map to obtain the final output results (for details, see Section 3.1).

Recent work has proven that self-supervised learning (Lin et al., 2017; Vaswani et al., 2017) can improve network performance. The SCA unit block is introduced in the network structure, as shown in Fig. 2. The SCA module re-calibrates the feature mapping between channels by self-supervising in combination with rich contextual information, emphasizes useful feature information, and suppresses useless feature information. The basic structure of SCA consists of two parts. The first part encodes the wider scene context information into local features, calculates the similarity between channels, and enhances its representation ability. The second part aggregates the spatial context information to improve the specific scene information in each channel and accurately adjusts the interdependence between the channels.

As shown in Fig. 2, firstly, input feature graph $X \in \mathbb{R}^{H\times W\times C}$ is generated by $\Phi(\cdot)$ and $\psi(\cdot)$ operations to generate new feature graphs $A \in \mathbb{R}^{H\times W\times C_1}$ and $B \in \mathbb{R}^{H\times W\times C_1}$ respectively. Among them, $\Phi(\cdot)$ and $\psi(\cdot)$ operations are convolutional layers including batch normalization (Ioffe and Szegedy, 2015) and ReLU layers (Nair and Hinton, 2010). In order to reduce the amount of calculation, set the size of the convolution kernel to $[1 \times 1 \times C_1]$, where $C_1 = \frac{1}{16}C$, reduce the channel dimension, and reduce the calculation amount of matrix multiplication. Then reshape the feature map A to $\mathbb{R}^{C_1\times HW}$ and transpose it to $\mathbb{R}^{HW\times C_1}$ after B reshape. Finally, multiply A and B in a matrix manner and use the softmax function to obtain a map $Z \in \mathbb{R}^{C_1\times C_1}$ of channel correlation. The calculation method is:

$$Z_{ij} = \frac{\exp\left(\psi(X_i)^T \cdot \Phi(X_j)\right)}{\sum_{j=1}^{HW} \exp\left(\psi(X_i) \cdot \Phi(X_j)\right)} \tag{1}$$

Where X_i represents the number i indexed pixel in the feature vector, j represents the index of all possible positions, and Z_{ij} represents the correlation between each remaining pixel and i.

At the same time, after the feature map X is input into $K(\cdot)$, the feature map $E \in \mathbb{R}^{H\times W\times C_1}$ is generated, and the feature map E reshape is transposed into $\mathbb{R}^{C_1\times HW}$. The operation of $K(\cdot)$ is the same as $\Phi(\cdot)$ and $\psi(\cdot)$. Then it is matrix-multiplied by the feature map Z to redistribute the correlation information to the original feature map. Then reshape the obtained result into $\mathbb{R}^{H\times W\times C_1}$ to obtain the feature map $D \in \mathbb{R}^{H\times W\times C_1}$. The calculation equation is as follows:

$$D_j = \sum_{j=1}^{HW} Z_{ij} K(X_i) \tag{2}$$

To construct the feature map D, the scene context mapping is aggregated according to the spatial attention mechanism, and the interrelated channels obtain mutual benefits (eq. 2).

In order to better adjust the correlation between each channel in the feature map X and other channels, the channel of the feature map D is weighted using the channel-wise module (CWB). First, the global average pooling is used to aggregate the spatial dimension $W \times H$ of the feature map D into a channel-wise statistic $\mathbf{v} \in \mathbb{R}^{C_1}$, where the number i

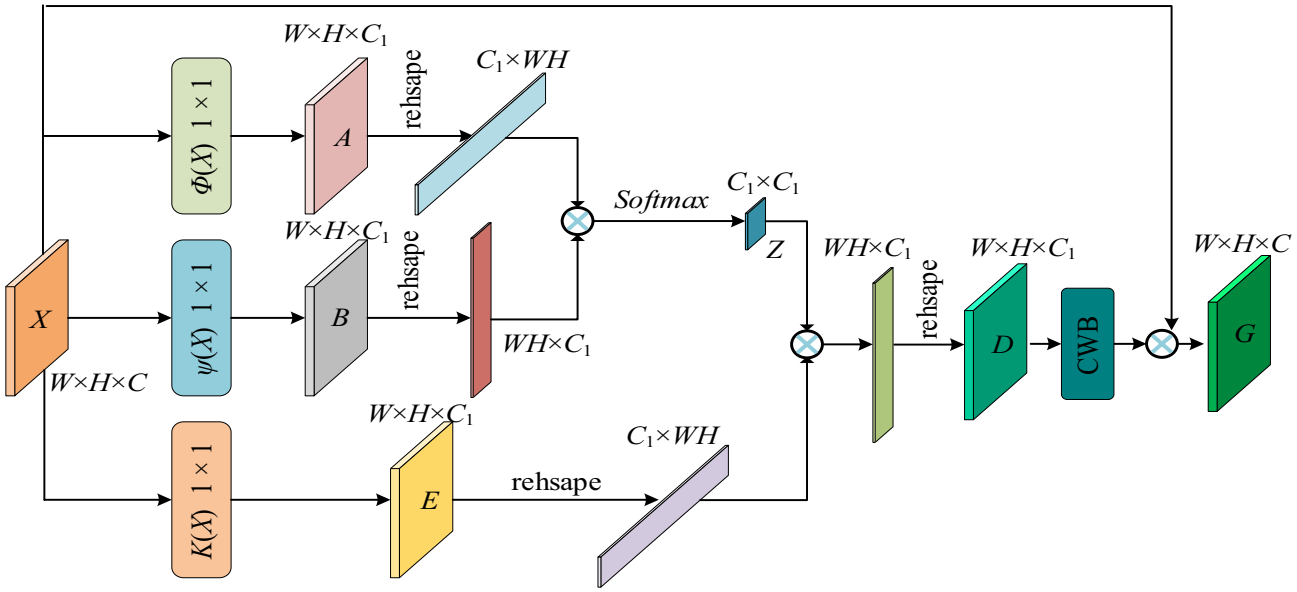


Fig. 2. Spatial and context attention block.

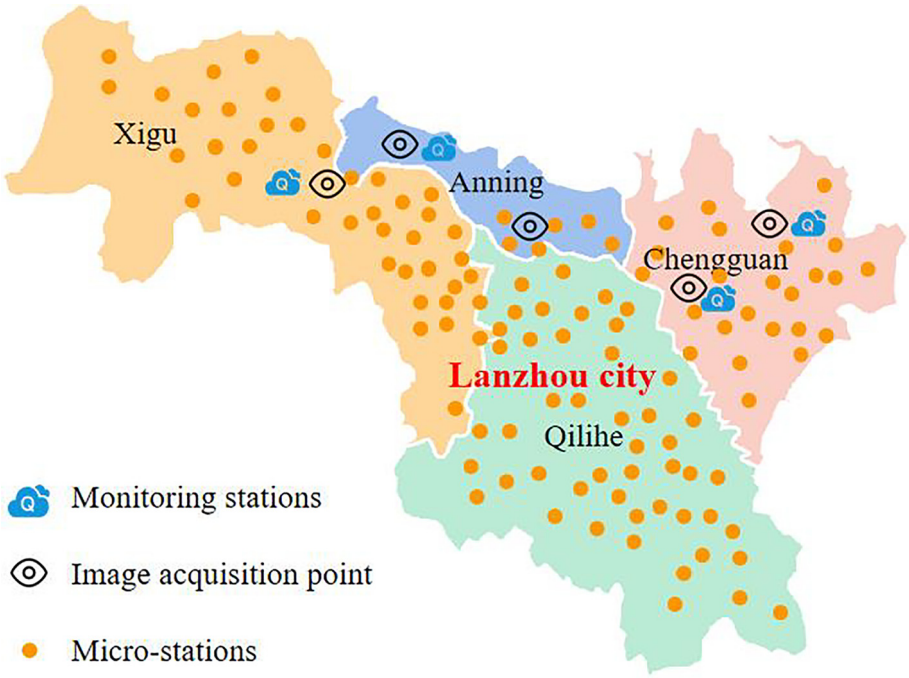


Fig. 3. The distribution of the monitoring stations.

element in **v** is calculated as follows:

$$v_i = \frac{1}{W \times H} \sum_{n=1}^{W} \sum_{m=1}^{H} D_i(n, m) \tag{3}$$

Feature map X contains C channels. A fully connected layer is added to adjust the dimension of **w** to $\mathbb{R}^C$. The calculation method is as follows:

$z = F(\mathbf{v}, \mathbf{W}) = \sigma(\mathbf{W}\mathbf{v})$, (4)where σ represents the Sigmoid activation function $\mathbf{W} \in \mathbb{R}^{C_1 \times C}$. The final output of the SCA module passes the readjusted feature map G:

$$g_i = F_{scale}(x_i, z_i) = z_i \cdot x_i \tag{5}$$

where $G = [g_1, g_2, g_n]$ and $F_{scale}(x_i, z_i)$ represents the feature map $X_i \in \mathbb{R}^{H \times W}$ multiplied by the weight z_i.

4. Experimental results and analysis

4.1. Model training

The model in this article is implemented using the pytorch (Paszke et al., 2017) deep learning framework. The server configuration for training the model is: Intel (R) Xeon (R) E5-2620 v3 2. 40GHz CPU, Tesla K80 GPU, and Ubuntu64 as OS. During the training process, the parameters are optimized using stochastic gradient descent (SGD), and the momentum β is set to 0. 9. To reduce the instability of the random gradient, mini-batch is set to 32. The initial learning rate is 10^{-2}, and it is reduced by 10 times in every 90 cycles, and the weight attenuation is 10^{-4}. The model uses the method in (He et al., 2015) to initialize the weights. All models start from 0 and are trained for 270 cycles.

4.2. Data set and preprocessing

There is no publicly available image library based on image detection for air quality, so a multi-scenario air quality image database (NWNU-AQI) was built. The library contains a total of 1241 images, which are a collection of scene images at different air quality levels. We took the following steps to build the dataset.

1) Set up a camera with a fixed position and orientation near Lanzhou's National Ambient Air Automatic Monitoring Station (monitoring station for short) or microsite to collect surrounding air quality images. The camera automatically captures sky images every hour. The layout of the collection points and monitoring stations is shown in Fig. 3. Table 1 details the air quality image collection point information. Fig. 4 is an example of scene images corresponding to different air pollution levels in the air quality image library. Source of air quality observation data. In recent years, more than 500 grid monitoring microstations (called micro-stations) for real-time monitoring of overall air quality have been deployed in Lanzhou. Each station can display six atmospheric pollutants monitored within a 1 km radius, including PM_{10} and $PM_{2.5}$, SO, NO, CO, O and AQI indicators.
2) Obtain the monitoring station data from the Chinese National Environmental Monitoring Station and micro-station data from the Lanzhou Environmental Protection Bureau.
3) Extract data of the base station closest to the camera and data of the micro stations located within a 1 km radius around the camera.

Table 1
Image acquisition information.

Numbering	Collection point	Photo pixels (Px)	Shooting time period	Collection interval
A	Railway Design Institute	4608 × 3456	7:00–19:00	One per hour
B	Biological Product Institute	3264 × 2448	7:00–19:00	One per hour
C	Education Port	3264 × 2448	7:00–19:00	One per hour
D	Lanlian hotel	4608 × 3456	7:00–19:00	One per hour
E	Northwest normal university	3264 × 2448	7:00–19:00	One per hour

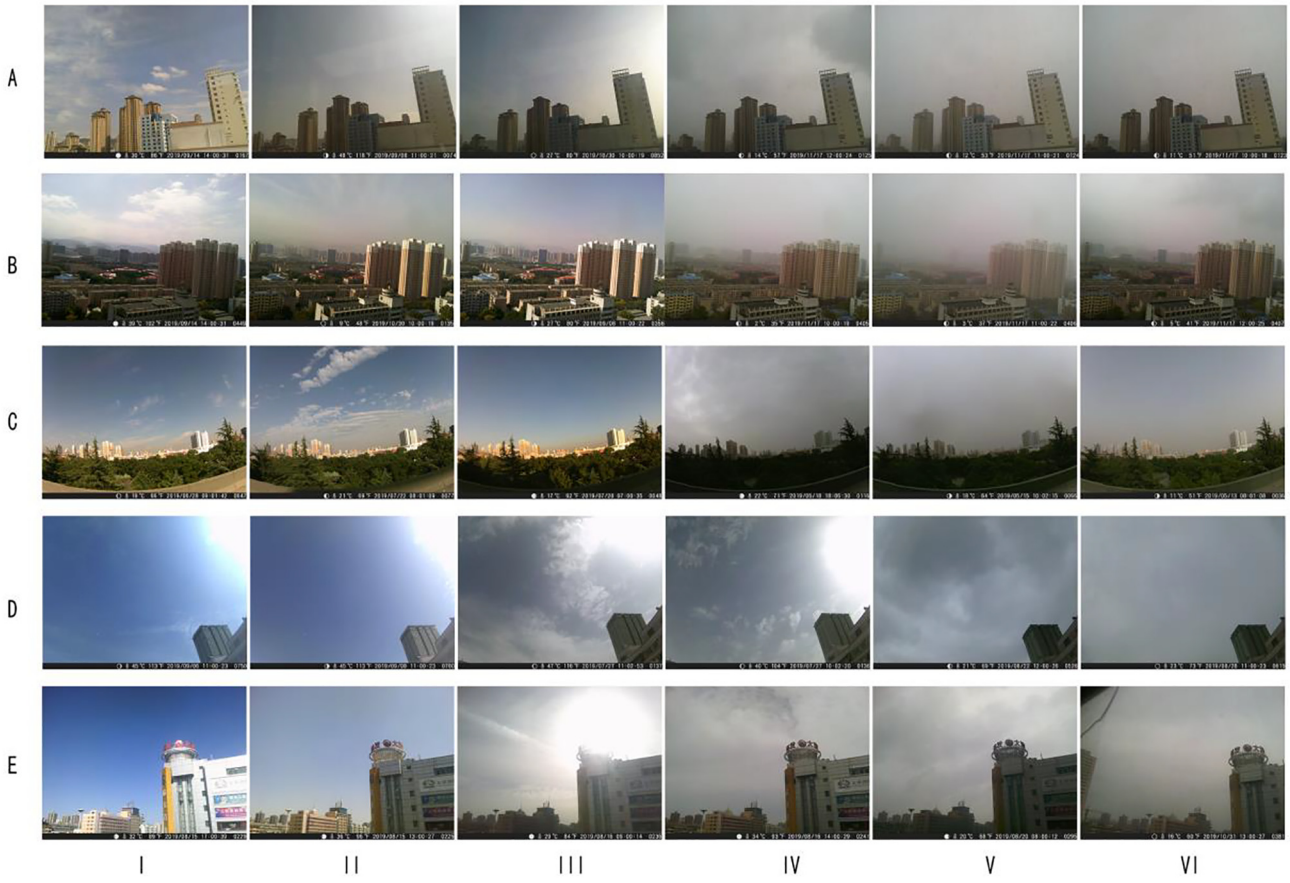


Fig. 4. Example images in NWNU-AQI.

Observation data of national base stations and micro-sites are hourly (one per hour), and the hourly values calculated by the AQI index are based on real-time observation data and non-daily average AQI. Station data and micro-site data were queried using timestamp correlation to obtain the national base station and micro-site observation table.

4) Obtain the average AQI with the highest reliability through weighted average of AQI, and convert to levels according to the AQI classification table.

5) Read the shooting time and file name of the images collected at each collection point in turn, and query the national base station and micro-site observation AQI data that match the nearby location based on the timestamp condition. It is formed by the shooting time, file name, AQI value, AQI level and other fields. In the dataset, there are a total of 1241 pieces of data, with the AQI level as the image label. In this way, the image acquisition and observation data that strictly correspond to the geographical location and time are formed, and a high-quality and same-site air quality image database is obtained (Table 2).

During the model training process, 70% of the images were randomly selected for the training set, and the remaining 30% used as the test set.

Due to the small number of samples in the training data set, and to prevent the model from overfitting and improve model accuracy and robustness, we enhanced the process of data set training as follows. Each training image is randomly sampled using the following methods: Randomly flip the image horizontally or vertically.

1) Rotate the image randomly between [0, 360°].
2) Randomly scale the image within a coefficient range of [0. 8, 1].
3) Crop the image to 3/4 or 4/3 of the original aspect ratio.

Table 2
Sample data set.

Shooting_time	Filename	Air_time	AQI	Grade	Primary pollutants
2019/1/6 8:00:00	DSCF0667.JPG	2019/1/6 8:00:00	113	3	$PM_{2.5}$
2019/1/6 9:00:00	DSCF0668.JPG	2019/1/6 9:00:00	115	3	$PM_{2.5}$
2019/1/6 10:00:00	DSCF0669.JPG	2019/1/6 10:00:00	127	3	$PM_{2.5}$
2019/1/6 11:00:00	DSCF0670.JPG	2019/1/6 11:00:00	128	3	$PM_{2.5}$
2019/1/6 12:00:00	DSCF0671.JPG	2019/1/6 12:00:00	122	3	$PM_{2.5}$
2019/1/6 13:00:00	DSCF0672.JPG	2019/1/6 13:00:00	102	3	$PM_{2.5}$
2019/1/6 14:00:00	DSCF0673.JPG	2019/1/6 14:00:00	84	2	$PM_{2.5}$
2019/1/6 15:00:00	DSCF0674.JPG	2019/1/6 15:00:00	58	2	PM_{10}
2019/1/6 16:00:00	DSCF0675.JPG	2019/1/6 16:00:00	55	2	PM_{10}
2019/1/6 17:00:00	DSCF0676.JPG	2019/1/6 17:00:00	55	2	PM_{10}

Table 3
The performance of the model with the SCA module on the NWNU-AQI test set.

Method	BaseNet	SCA	P-Times(s)	ACCmean(%)
ResNet	ResNet18	N	0.0033	68.2
ResNet-Data Augmentation	ResNet18	N	0.0033	70.1
AQC-Net(Ours)	ResNet18	Y	0.0042	69.3
AQC-Net(Ours)-Data Augmentation	ResNet18	Y	0.0042	74.0

After the above steps, the size of each sampling area is adjusted to 256 × 256, and then normalized to [0, 1].

4.3. Spatial and context attention block with network

In this set of experiments, the performance of ResNet18 and AQC-Net on the data set NWNU-AQI and the performance of the model before and after the data enhancement were compared. Different models were trained using the same optimization strategy. The results of different models on the test set are shown in Table 3. The SCA module can significantly improve the performance of the model. The classification accuracy of AQC-Net on the test set is 74%. Compared with 70.1% of ResNet18, the introduction of the SCA module can improve the accuracy by 1. 9%. In terms of response time, the channel in the SCA module was reduced in dimension before matrix multiplication, which reduced the complexity and computation of the model. Comparing the two models, ResNet and AQC-Net, for an input picture of 256 × 256 pixels, the time required for ResNet to make a forward pass is about 0.0038 s. The feature extraction and matrix calculation of the SCA module through several convolutional layers will increase the time consumption. In general, the time for one pass of AQC-Net is about 0. 0042 s.

Fig. 5 shows the confusion matrix for model classification accuracy: Fig. 5(a) shows the predicted confusion matrix, and Fig. 5(b) shows the standardized confusion matrix. Horizontal 1 to 6 represent the predicted values of the test samples, and vertical 1 to 6 represent the true values of the test samples. The number of model predictions consistent with the true values of the test samples falls on the diagonal of the confusion matrix. On-diagonal values in Fig. 5(a) indicate the number of correctly classified images, and off-diagonal values indicate the number of misclassified images that deviates from the diagonal. For example, 82 test samples at AQI level 1 are correctly classified, and 36 test samples at AQI level 2 are incorrectly identified. On-diagonal values in Fig. 5 (b) indicate the proportion of accurate predictions. AQI level 2 is 31%, for example.

To further verify the method in this paper, the class activation diagrams (Zhou et al., 2016) generated by ResNet and AQC-Net are compared in Fig. 6. The image samples are randomly selected from the test set in NWNU-AQI, where the first line is the result produced by ResNet and the second line is the result produced by AQC-Net. In (a) (b) samples, ResNet prediction is wrong, and (c) (d) (e) (f) prediction is correct. But AQC-Net made correct predictions for all six samples. As shown in the figure, in the NWNU-AQI data set, AQC-Net tends to focus on a wider area, and has more contextual feature information combined. The area of interest for ResNet is relatively small, and more comprehensive contextual and detailed information cannot be obtained.

4.4. Model performance comparison

The traditional machine learning methods SVM (Adankon and Cheriet, 2009) and deep learning methods VGG (Liu and Deng, 2015), ResNet (Vaswani et al., 2017) and AQC-Net were compared and analyzed on the NWNU-AQI dataset. The core idea of the SVM classifier is to transform image classification problems into high-dimensional feature classification spaces, and difficult-to-classify problems to be linearly separable. In the high-dimensional feature classification space, a kernel is used to generate a hyperplane to distinguish different air quality levels. Because the image classification problem has linear inseparability, this paper uses RBF radial basis kernel. The accuracy of SVM model detection is 60%, and it is time-consuming. The average time to classify a picture is 0. 0435 s. VGG improves the accuracy of the model by deepening the depth of the network, using small convolution kernels instead of large ones, and achieves better results in air level detection. The accuracy of VGG for air level detection is 68. 3%, which is 8. 3% higher than SVM, and it is less time consuming. The idea of using the residual module in ResNet solves the problem of gradient disappearance when deepening the network model, and is more sensitive to changes in model weights. The accuracy of using ResNet18 to detect air quality levels increases by 1. 8% on VGG16, and the speed increases by nearly 2 times. The AQC-Net proposed in this paper, by adding the SCA module to the third block of ResNet18, improves the accuracy of model detection to 74% (Table 4).

Since each channel of the feature map output by the high-level convolutional layer can be regarded as a response of a specific class, and is dependent on each other, We introduced the SCA module in the high-level convolutional layer. By calculating the direct similarity between channels, the interdependence between these channel maps is used to dynamically adjust the interdependence between feature

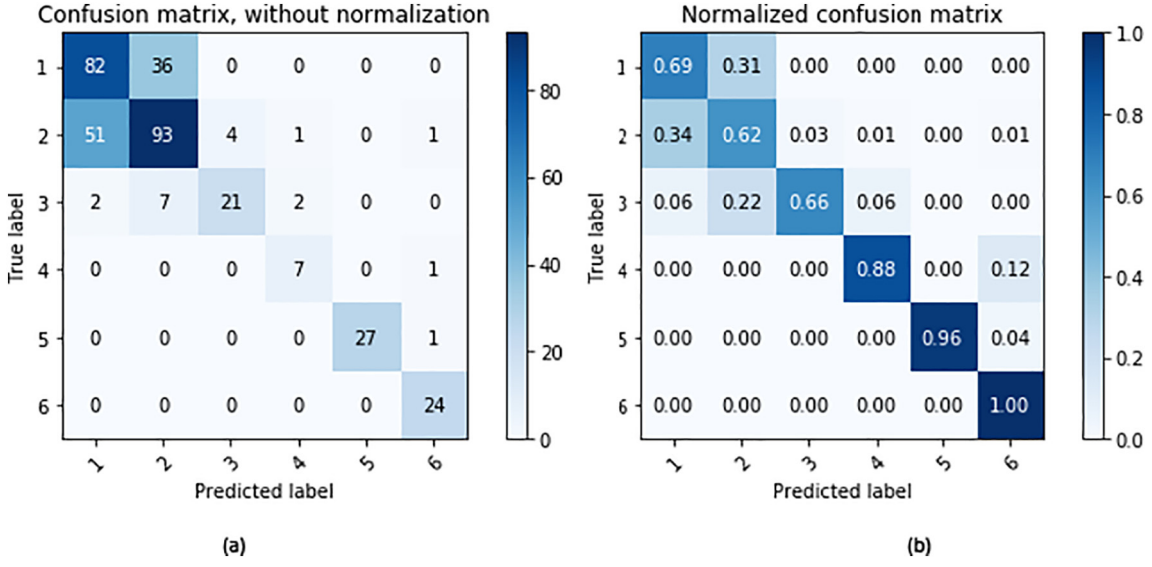


Fig. 5. Confusion matrix.

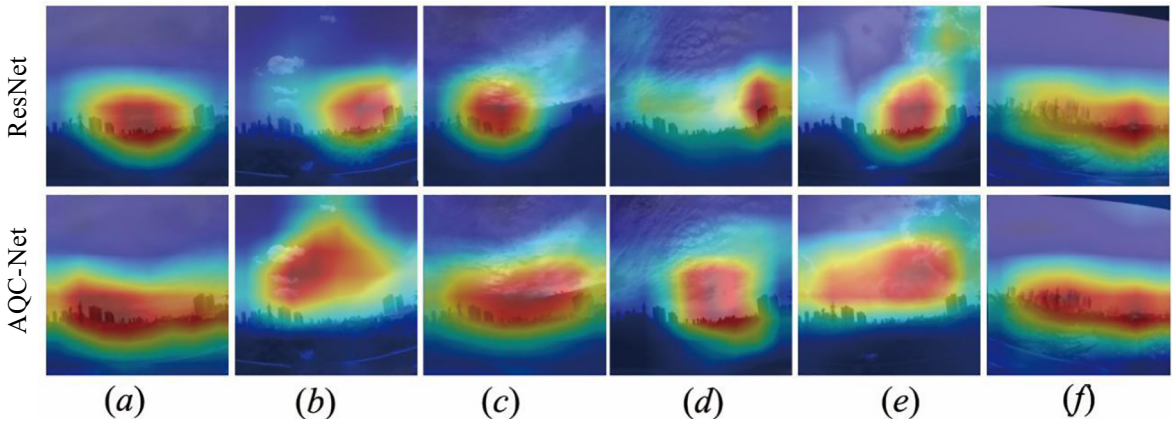


Fig. 6. Comparison of ResNet and AQC-Net activation diagram.

maps to improve the feature map. The ability to represent allows the model to detect air quality levels more accurately.

In addition, we used 5 indicators to evaluate the predictive performance of SVM, VGG, ResNet, and AQC-Net. They are Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Absolute Percentage Error (SMAPE). The evaluation results are shown in Table 5. AQC-Net outperforms other models across all these indicators.

Fig. 7 is the comparison between the observed and predicted values of the four models in the time series. The horizontal axis is the acquisition time of the image and the vertical axis is the AQI value.

Fig. 8 is a scatter plot of the observed and predicted values of the four models. The horizontal axis is the serial number of the test sample and the vertical axis is the value of AQI.

5. Discussion and limitations

Because the data set used in this article consists of fixed-point air quality image only, the angle is facing the sky and about 1/3 of the picture is buildings and ground. The intention is to simulate a more common and easier shooting angle. For monitoring, 50% or more of the frames shot are in the sky. The images used in this experiment all reflect daytime situations (between 7:00 and 19:00). Due to the poor imaging quality at night, the visibility is very low. This experimental model only considers air quality observation during the day and is not suitable for night time. Because the model training data is collected in Lanzhou, the model has relatively high controllability, reliability and efficiency in the local area and the model's prediction speed and accuracy are relatively stable. But due to different regional climate and atmospheric differences, the model may not be able to achieve the required accuracy in other regions. To use our model in other regions, it needs to be trained and adjusted again using local image data.

This model cannot reach the monitoring accuracy of the air quality monitoring stations due to various constraints, but can play a complementary role. The advantage of the model is that people can use portable image acquisition equipment to obtain comprehensive air quality indicators at the location in real time, especially for areas that are far away from the monitoring sites, such as suburban or rural areas.

Several areas can be improved in future research. Different weather conditions mainly affect the lightness or darkness of air quality images. The brightness characteristics of the image can be directly extracted from the model. Likewise, humidity does not have a significant effect on air quality images, although it may affect air quality. Future research can incorporate these factors to improve model accuracy. Finally, our research focused on the AQI, which is a comprehensive indicator of air quality. Future research, if interested, can also focus specifically on $PM_{2.5}$.

6. Conclusion

This paper proposes a deep learning-based air quality rating model AQC-Net. This model mainly uses deep convolutional neural networks to extract feature representation information related to air quality in scene images to classify air quality levels. In this paper, the self-supervision mechanism, SCA module, is introduced into the model to automatically capture the interdependence between the air quality information in the scene image, and strengthen or suppress the local feature information between the channels to improve the model's representation ability. The experimental results show that the SCA module can effectively improve the accuracy of model classification and is more suitable for air quality rating assessment than other methods. In our future work, how to further improve the accuracy of the model will be studied.

CRediT authorship contribution statement

Qiang Zhang: Conceptualization, Data curation, Formal analysis, Investigation, Writing - review & editing, Supervision.**Fengchen Fu:**Validation, Formal analysis, Visualization, Writing - review & editing, Data curation.**Ran Tian:**Resources, Writing - review & editing, Supervision.

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.

Table 4
Performance of different methods on the NWNU-AQI test set.

Method	BaseNet	Times(s)	ACCmean(%)
SVM (Adankon and Cheriet, 2009)	–	0.0435	60.0
VGG (Liu and Deng, 2015)	VGG16	0.0073	68.3
ResNet (He et al., 2016)	ResNet18	0.0038	70.1
AQC-Net (Ours)	ResNet18	0.0042	74.0

Table 5
Indicator values of SVM, VGG, ResNet and AQC-Net.

Indicator	SVM	VGG	ResNet	AQC-Net
MSE	0.4972	0.4010	0.4213	0.3426
RMSE	0.7051	0.6333	0.6491	0.5854
MAE	0.4194	0.3299	0.3401	0.2919
MAPE	23.1806	21.2521	22.4788	18.6633
SMAPE	20.4146	18.9785	19.3339	17.3993

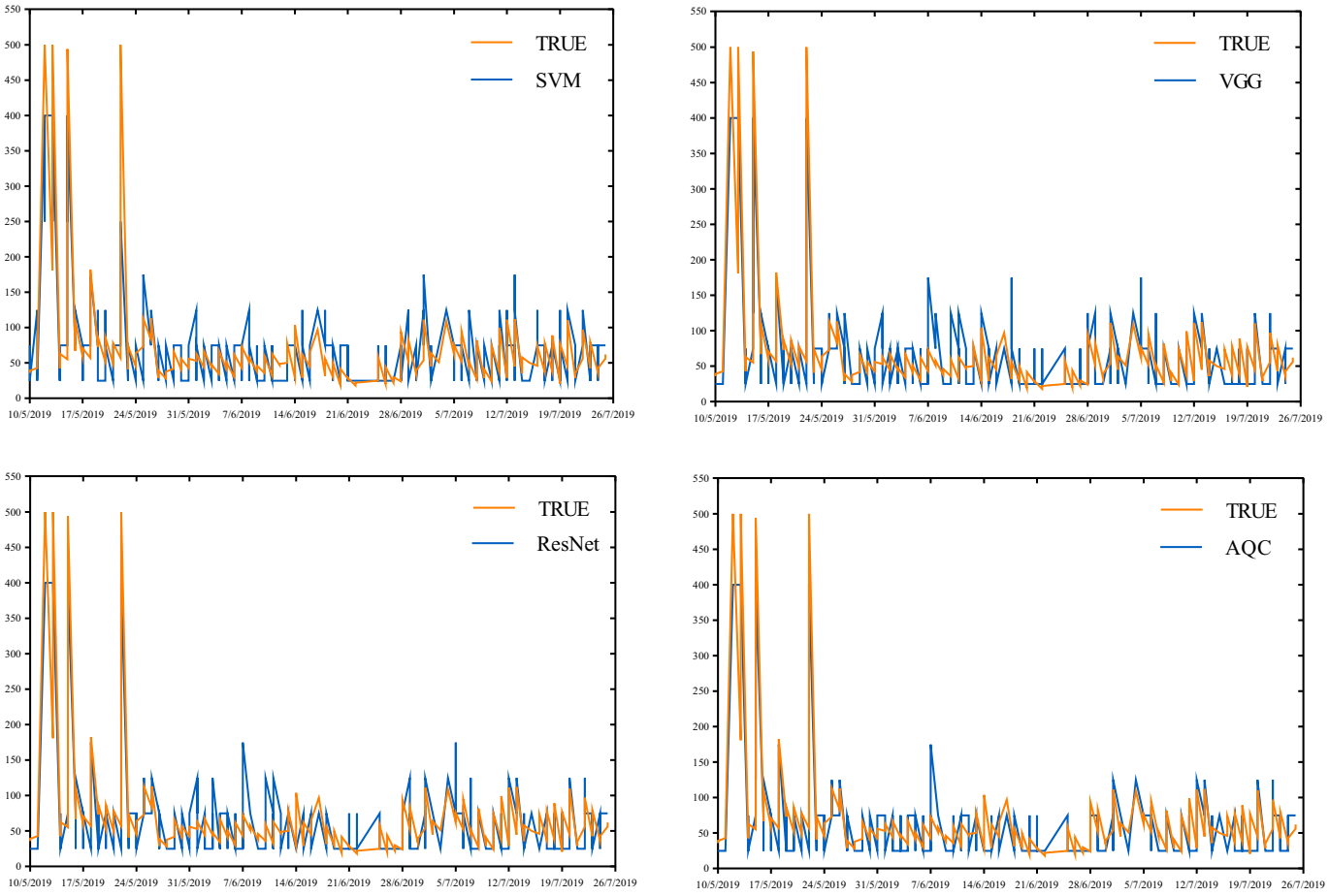


Fig. 7. Time series of the observed values and the predicted values of the models.

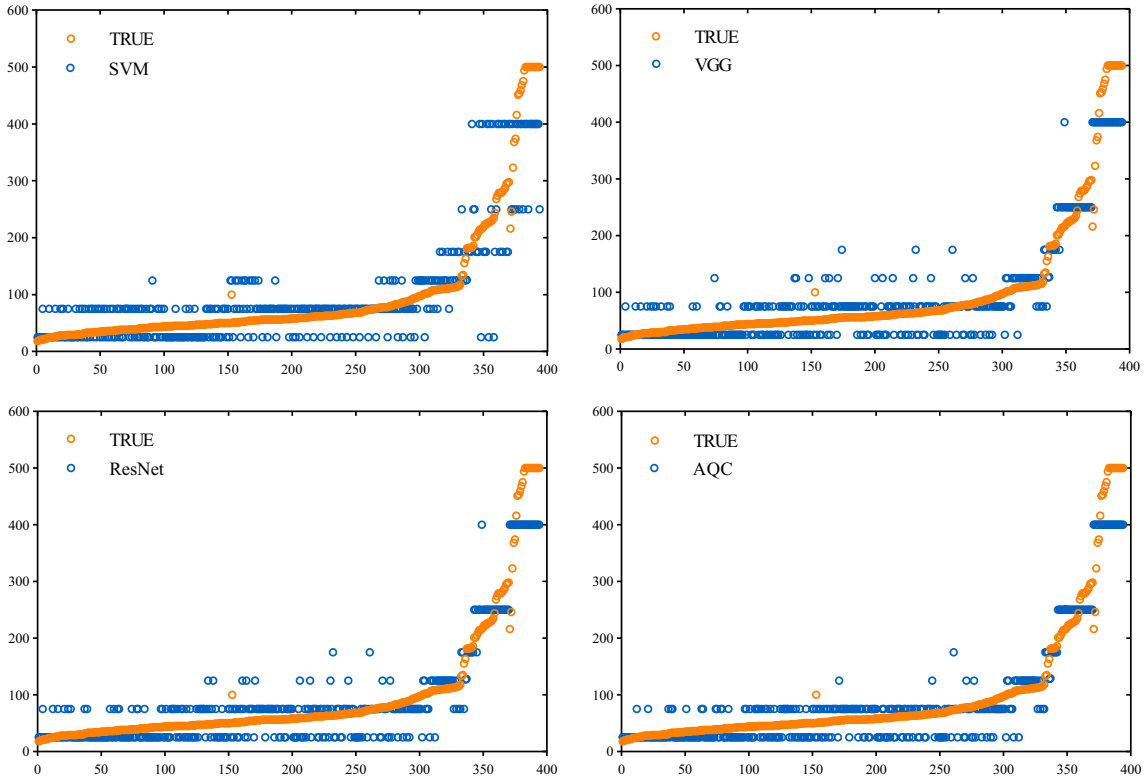


Fig. 8. Scatter diagram of the observed values and the predicted values of the models.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 71764025 and No. 71961028).

References

Adams, M.D., Kanaroglou, P.S., 2016. Mapping real-time air pollution health risk for environmental management: combining mobile and stationary air pollution monitoring with neural network models. J. Environ. Manag. 168, 133–141. https://doi.org/10.1016/j.jenvman.2015.12.012.

Adankon, M.M., Cheriet, M., 2009. Support Vector Machine. Encyclopedia of Biometrics. , pp. 1303–1308. https://doi.org/10.1109/TNN.2009.2031143.

Alson, J.N., Misagal, L.V., 2016. Smart phones usage among college students. IMPACT: International Journal of Research in Engineering & Technology (IMPACT: IJRET) 4, 63–70.

Ameer, S., Shah, M.A., Khan, A., Song, H., Maple, C., Islam, S.U., et al., 2019. Comparative analysis of machine learning techniques for predicting air quality in smart cities. IEEE Access 7, 128325–128338. https://doi.org/10.1109/ACCESS.2019.2925082.

Bo, Q., Yang, W., Rijal, N., Xie, Y., Feng, J., Zhang, J., 2018. Particle pollution estimation from images using convolutional neural network and weather features. 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, pp. 3433–3437. https://doi.org/10.1109/ICIP.2018.8451306.

Burnett, R.T., Pope III, C.A., Ezzati, M., Olives, C., Lim, S.S., Mehta, S., et al., 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ. Health Perspect. 122, 397–403. https://doi.org/10.1289/ehp.1307049.

Chakma, A., Vizena, B., Cao, T., Lin, J., Zhang, J., 2017. Image-based air quality analysis using deep convolutional neural network. 2017 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 3949–3952. https://doi.org/10.1109/ICIP.2017.8297023.

Chen, J., Chen, H., Zheng, G., Pan, J.Z., Wu, H., Zhang, N., 2014. Big smog meets web science: smog disaster analysis based on social media and device data on the web. Proceedings of the 23rd International Conference on World Wide Web. ACM, pp. 505–510. https://doi.org/10.1145/2567948.2576941.

Chen X, Li Y, Li D. An efficient method for air quality evaluation via ANN-based image recognition. 2016 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE 2016). Atlantis Press, 2016. https://dx.doi.org/https://doi.org/10.2991/aiie-16.2016.59.

Chestnut, L.G., Schwartz, J., Savitz, D.A., Burchfiel, C.M., 1991. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. Arch. Environ. Health 46, 135–144. https://doi.org/10.1080/00039896.1991.9937440.

Chipps, B.E., 2015. Association of Improved air Quality with Lung Development in children. Pediatrics 136, S233. https://doi.org/10.1056/NEJMoa1414123.

Chung, Y., 1986. Air pollution detection by satellites: the transport and deposition of air pollutants over oceans. Atmos. Environ. 20, 617–630. https://doi.org/10.1016/0004-6981(86)90177-0.

De-zhong, D., 2013. Advance on comparison research of monitoring techniques for PM2. 5 in ambient air. China Test 39, 1–5. http://www.cqvip.com/qk/93607c/201302/46620646.html.

Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., et al., 1993. An association between air pollution and mortality in six US cities. N. Engl. J. Med. 329, 1753–1759. https://doi.org/10.1056/NEJM199312093292401.

Fattal, R., 2008. Single image dehazing. ACM Trans. Graph. 27, 72.

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, 5880–5892. https://doi.org/10.1016/j.atmosenv.2006.03.016.

Hauck, H., Berner, A., Gomiscek, B., Stopper, S., Puxbaum, H., Kundi, M., et al., 2004. On the equivalence of gravimetric PM data with TEOM and beta-attenuation measurements. J. Aerosol Sci. 35, 1135–1149. https://doi.org/10.1016/j.jaerosci.2004.04.004.

He, K., Sun, J., Tang, X., 2010. Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell. 33, 2341–2353. https://doi.org/10.1109/TPAMI.2010.168.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. Proceedings of the IEEE international conference on computer vision, 1026–1034 https://doi.org/10.1109/ICCV.2015.123.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 770–778 https://doi.org/10.1109/CVPR.2016.90.

Hodgeson, J., McClenny, W., Hanst, P., 1973. Air pollution monitoring by advanced spectroscopic techniques: a variety of spectroscopic methods are being used to detect air pollutants in the gas phase. Science 182, 248–258. https://doi.org/10.1126/science.182.4109.248.

Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint arXiv:1502.03167. https://arxiv.org/abs/1502.03167.

Knollenberg, R.G., 1970. The optical array: an alternative to scattering or extinction for airborne particle size determination. J. Appl. Meteorol. 9, 86–103. https://doi.org/10.1175/1520-0450(1970)009%3C0086:TOAAAT%3E2.0.CO;2.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, pp. 1097–1105. https://doi.org/10.1145/3065386.

Li, L., C-z, Z.H.A.I., J-y, Y.U., 2012. A Review of Domestic and Overseas Research on Air Quality Monitoring Networks Designing. Environmental Monitoring in China. 13. http://en.cnki.com.cn/Article_en/CJFDTotal-IAOB201204013.htm.

Li, L., Zheng, Y., Zhang, L., 2014. Demonstration abstract: PiMi air box: A cost-effective sensor for participatory indoor quality monitoring. Proceedings of the 13th International Symposium on Information Processing in Sensor Networks. IEEE Press, pp. 327–328. https://doi.org/10.1109/IPSN.2014.6846786.

Li, F., Liu, Y., Lü, J., Liang, L., Harmer, P., 2015a. Ambient air pollution in China poses a multifaceted health threat to outdoor physical activity. J. Epidemiol. Community Health 69, 201–204. https://doi.org/10.1136/jech-2014-203892.

Li, Y., Huang, J., Luo, J., 2015b. Using user generated online photos to estimate and monitor air pollution in major cities. Proceedings of the 7th International Conference on Internet Multimedia Computing and Service. ACM, p. 79. https://doi.org/10.1145/2808492.2808564.

Li, X., Peng, L., Hu, Y., Shao, J., Chi, T., 2016. Deep learning architecture for air quality predictions. Environ. Sci. Pollut. Res. 23, 22408–22417. https://doi.org/10.1007/s11356-016-7812-9.

Lin, M., Chen, Q., Yan, S., 2013. Network in Network. arXiv preprint arXiv:1312.4400. https://arxiv.org/abs/1312.4400.

Lin, Z., Feng, M., CNd, Santos, Yu, M., Xiang, B., Zhou, B., et al., 2017. A Structured Self-attentive Sentence Embedding. arXiv preprint arXiv:1703.03130. https://arxiv.org/abs/1703.03130.

Liu, S., Deng, W., 2015. Very deep convolutional neural network based image classification using small training sample size. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). IEEE, pp. 730–734. https://doi.org/10.1109/ACPR.2015.7486599.

Liu, H.T, Li, F., Li, F.G, Li, H.Q, 2011. The evaluation of air quality using image quality. Journal Of Image And Graphics 16, 1030–1037. http://www.cnki.com.cn/Article/CJFDTotal-ZGTB201106020.htm.

Liu, P., Zheng, J., Li, Z., Zhong, L., Wang, X., 2010. Optimization of site locations of regional air quality monitoring network: methodology study. China Environ. Sci. 30, 907–913. https://doi.org/10.3724/SP.J.1088.2010.00432.

Liu, F., Shen, C., Lin, G., Reid, I., 2015a. Learning depth from single monocular images using deep convolutional neural fields. IEEE Trans. Pattern Anal. Mach. Intell. 38, 2024–2039. https://doi.org/10.1109/TPAMI.2015.2505283.

Liu, X., Song, Z., Ngai, E., Ma, J., Wang, W., 2015b. PM2: 5 monitoring using images from smartphones in participatory sensing. 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, pp. 630–635. https://doi.org/10.1109/INFCOMW.2015.7179456.

Liu, C., Tsow, F., Zou, Y., Tao, N., 2016. Particle pollution estimation based on image analysis. PLoS One 11, e0145955. https://doi.org/10.1371/journal.pone.0145955.

Ma, J., Li, K., Han, Y., Yang, J., 2018. Image-based air pollution estimation using hybrid convolutional neural network. 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, pp. 471–476. https://doi.org/10.1109/ICPR.2018.8546004.

Mao, J., Phommasak, U., Watanabe, S., Shioya, H., 2014. Detecting foggy images and estimating the haze degree factor. Journal of Computer Science & Systems Biology 7, 226–228. https://doi.org/10.4172/jcsb.1000161.

Mei, S., Li, H., Fan, J., Zhu, X., Dyer, C.R., 2014. Inferring air pollution by sniffing social media. Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE Press, pp. 534–539. https://doi.org/10.1109/ASONAM.2014.6921638.

Muir, D., Laxen, D.P., 1995. Black smoke as a surrogate for PM10 in health studies? Atmos. Environ. 29, 959–962. https://doi.org/10.1016/1352-2310(94)00370-Z.

Murty, R.N., Mainland, G., Rose, I., Chowdhury, A.R., Gosain, A., Bers, J., et al., 2008. Citysense: an urban-scale wireless sensor network and testbed. 2008 IEEE conference on technologies for homeland security. IEEE, 583–588 https://doi.org/10.1109/THS.2008.4534518.

Nader, J.S., 1975. Current technology for continuous monitoring of particulate emissions. J. Air Pollut. Control Assoc. 25, 814–821. https://doi.org/10.1080/00022470.1975.10470141.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10) , pp. 807–814. https://www.cs.toronto.edu/~hinton/absps/reluICML.pdf.

Padayachi, Y.R., 2016. Satellite Remote Sensing of Particulate Matter and Air Quality Assessment in the Western Cape, South Africa. http://ukzn-dspace.ukzn.ac.za/handle/10413/15488.

Pan, Z., Yu, H., Miao, C., Leung, C., 2017. Crowdsensing air quality with camera-enabled mobile devices. Twenty-Ninth IAAI Conference https://www.aaai.org/ocs/index.php/IAAI/IAAI17/paper/viewPaper/14171.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al., 2017. Automatic Differentiation in Pytorch. https://openreview.net/forum?id=BJJsrmfCZ.

Patashnick, H., Rupprecht, E.G., 1991. Continuous PM-10 measurements using the tapered element oscillating microbalance. J. Air Waste Manage. Assoc. 41, 1079–1083. https://doi.org/10.1080/10473289.1991.10466903.

Portney, P.R., Mullahy, J., 1990. Urban air quality and chronic respiratory disease. Reg. Sci. Urban Econ. 20, 407–418. https://doi.org/10.1016/0166-0462(90)90019-Y.

Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., et al., 2013. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European study of cohorts for air pollution effects (ESCAPE). Lancet Oncol. 14, 813–822. https://doi.org/10.1016/S1470-2045(13)70279-1.

Rijal, N., Gutta, R.T., Cao, T., Lin, J., Bo, Q., Zhang, J., 2018. Ensemble of deep neural networks for estimating particulate matter from images. 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). IEEE, pp. 733–738. https://doi.org/10.1109/ICIVC.2018.8492790.

Rui, L., Ming-Shun, Z., 2015. Air Quality Monitoring in European Union and Recommendations for China Strengthening Air Quality Monitoring System. Environmental Monitoring Management and Technology. , pp. 7–10. http://www.cnki.com.cn/Article/CJFDTotal-HJJS201502003.htm.

Schwartz, J., 1989. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. Environ. Res. 50, 309–321. https://doi.org/10.1016/S0013-9351(89)80012-X.

Smith, J.D., Atkinson, D.B., 2001. A portable pulsed cavity ring-down transmissometer for measurement of the optical extinction of the atmospheric aerosol. Analyst 126, 1216–1220. https://doi.org/10.1039/B101491I.

Tan, R.T., 2008. Visibility in bad weather from a single image. 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 1–8. https://doi.org/10.1109/CVPR.2008.4587643.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al., 2017. Attention is all you need. Advances in Neural Information Processing Systems , pp. 5998–6008. https://arxiv.org/abs/1706.03762.

Wang, H., Yuan, X., Wang, X., Zhang, Y., Dai, Q., 2014. Real-time air quality estimation based on color image processing. 2014 IEEE Visual Communications and Image Processing Conference. IEEE, pp. 326–329. https://doi.org/10.1109/VCIP.2014.7051572.

Wen-wen, Y., 2019. Research on Image Air Quality Evaluation Algorithm Based on Convolutional Neural Network. Northwest University http://cdmd.cnki.com.cn/Article/CDMD-10697-1019663758.htm.

Yong-hong, L., 2011. Characteristic analysis on uneven distribution of air pollution in cities [J]. Environmental Monitoring in China 3. http://en.cnki.com.cn/Article_en/CJFDTotal-IAOB201103024.htm.

Younan, D., Petkus, A.J., Widaman, K.F., Wang, X., Casanova, R., Espeland, M.A., et al., 2019. Particulate matter and episodic memory decline mediated by early neuroanatomic biomarkers of Alzheimer's disease. Brain https://doi.org/10.1093/brain/awaa007.

Cui, Y.-Q., Wang, C.-Y., Shang, Y.-C., 2012. Development of particulate matter in ambient air (PM_ (2.5)) monitoring technique [J]. China Environmental Protection Industry 4 (doi:CNKI:SUN:ZHBY.0.2012-04-012).

Yu, X., Liu, Y., Zhu, Y., Feng, W., Zhang, L., Rashvand, H.F., et al., 2012. Efficient sampling and compressive sensing for urban monitoring vehicular sensor networks. IET Wireless Sens. Syst. 2, 214–221. https://doi.org/10.1049/iet-wss.2011.0121.

Yu, J., Huang, B., Cheng, X., Atiquzzaman, M., 2016. Shortest link scheduling algorithms in wireless networks under the SINR model. IEEE Trans. Veh. Technol. 66, 2643–2657. https://doi.org/10.1109/TVT.2016.2580379.

Zhan, D., Kwan, M.-P., Zhang, W., Yu, X., Meng, B., Liu, Q., 2018. The driving factors of air quality index in China. J. Clean. Prod. 197, 1342–1351. https://doi.org/10.1016/j.jclepro.2018.06.108.

Zhang, Z., Ma, H., Fu, H., Wang, X., 2015. Outdoor air quality inference from single image. International Conference on Multimedia Modeling. Springer, pp. 13–25. https://doi.org/10.1007/978-3-319-14442-9_2.

Zhang, C., Yan, J., Li, C., Rui, X., Liu, L., Bie, R., 2016a. On estimating air pollution from photos using convolutional neural network. Proceedings of the 24th ACM International Conference on Multimedia. ACM, pp. 297–301. https://doi.org/10.1145/2964284.2967230.

Zheng, Y., Liu, F., Hsieh, H.-P., 2013. U-air: When urban air quality inference meets big data. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1436–1444. https://doi.org/10.1145/2487575.2488188.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016. Learning deep features for discriminative localization. Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2921–2929 https://doi.org/10.1109/CVPR.2016.319.