

Received May 6, 2019, accepted May 27, 2019, date of publication June 5, 2019, date of current version June 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2920865

Weather Visibility Prediction Based on Multimodal Fusion

CHUANG ZHANG¹, (Member, IEEE), MING WU¹, JINYU CHEN¹, KAIYAN CHEN¹,
CHI ZHANG², CHAO XIE³, BIN HUANG³, AND ZICHEN HE⁴

¹Pattern Recognition and Intelligent System Lab, Beijing University of Posts and Telecommunications, Beijing 100876, China

²Division of Energy Processes, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

³National Meteorological Center, Beijing 100081, China

⁴Department of Statistics and Computer Science, McGill University, Montreal, QC H3A0G4, Canada

Corresponding authors: Ming Wu (wuming@bupt.edu.cn) and Chuang Zhang (zhangchuang@bupt.edu.cn)

This work was supported by the National Natural Science Fund under Grant 61773071.

ABSTRACT Visibility affects all forms of traffic: roads, sailing, and aviation. Visibility prediction is meaningful in guiding production and life. Different from weather prediction, which relies solely on atmosphere factors, the factors that affect meteorological visibility are more complicated, such as the air pollution caused by factory exhaust emission. However, the current prediction of visibility is mostly based on the numerical prediction method similar to the weather prediction. We proposed a method using multimodal fusion to build a weather visibility prediction system in this paper. An advanced numerical prediction model and a method for emission detection were used to build a multimodal fusion visibility prediction system. We used the most advanced regression algorithm, XGBoost, and LightGBM, to train the fusion model for numerical prediction. Through the estimation of factory emission by the traditional detector in the satellite image, we propose to add the result of estimation based on Landsat-8 satellite images to assist the prediction. By testing our numerical model in atmosphere data of various meteorological observation stations in Beijing–Tianjin–Hebei region from 2002 to 2018, our numerical prediction model turns out to be more accurate than other existing methods, and after fusing with emission detection method, the accuracy of our visibility prediction system has been further improved.

INDEX TERMS Visibility prediction, emission estimation, numerical prediction, satellite image.

I. INTRODUCTION

Visibility is the farthest distance that a person with normal vision can recognize a target from the background. Visibility information is essential in air quality monitoring which needs to accurate real-time observations of visibility. Severe reduction of visibility often results in great inconveniences to our production and life, and even major traffic accidents. Thus, it is of great significance to establish an accurate forecast visibility system. To satisfy the needs for accurately measuring visibility, many models have emerged to solve the problem [1]. The works in [2], [3] indicated that the application on real-time data can generalize on the predicted data. So it is reasonable to use the real-time data to research the correlation between visibility and possible causes. Thus we focus on exploring a viable solution to build a highly accurate visibility prediction system.

Traditional methods [4], [5] of visibility prediction focus on the correlation between meteorological factors and

visibility, the construction of the traditional forecasting system relies on meteorological knowledge and subjective experience [6]–[8]. With the rise of machine learning, many visibility prediction schemes using machine learning have been proposed in recent years [2], [3], [9], [10]. These methods use a large number of meteorological data to search the latent relation between the input and visibility, and then applied it to visibility prediction.

The data used in the current works are limited to the meteorological field, and some spatial and temporal information have difficulties in characterizing by numerical data, resulting in the less accurate performance of the model.

Because of the aforementioned deficiency, we propose some strategies to improve the representation of the meteorological data. Feature engineering and model fusion based on distribution have been introduced to improve the model of numerical visibility prediction.

Through observation of such fact, by the mandatory temporary closure of factory exhaust emissions, NO_x and VOCs were reduced by 202 and 493 tons respectively by stopping

or limiting the major industries during APEC, and the local contribution of these measures to the decline in PM_{2.5} during the meeting was 17.5 percentage. Therefore, the correlation between exhaust emissions and visibility prediction can be proved by the fact. However, due to incomplete statistics on factories, the prediction of weather visibility cannot be entirely dependent on the estimation of factory emissions.

It is natural to think of fusing the information of the factory emission with numerical prediction of visibility. Such thought was popularly used as the multimodal fusion.

Many of the providers of multimodal data are associated with high impact commercial, social, biomedical, environmental, and military applications, and thus motivate the development of new and efficient analytical methodologies. It utilizes the complementarity between multimodal states to remove the redundancy between modalities to improve the performance of the model. Since satellite imagery as a rapidly evolving technology can intuitively and accurately reflect the real situation of the ground. Although the available satellite image is the real-time data, by further introducing satellite images' deduction we can also obtain the predicted images to implement the visibility prediction system. By taking into account the spatial correlation between factory emissions and visibility, satellite images can be used to estimate the emissions of the studied region. Thus, the factories' emission information has been introduced in our model as the second modal information to assist the prediction of the visibility. As a result, a visibility prediction scheme based on multimodal learning consisted of numerical prediction results and the estimation of satellite image-based exhaust emission has been proposed.

In this paper, the proposed method contributes to improving the numerical prediction by using the strategies to augment the feature of meteorological data, we name the system as MFVP (Multimodal Fusing for Visibility Prediction). Innovatively, by introducing the multimodal information of the estimation of factory emission based on satellite images, the result of prediction turns out to surpass the state of the art methods' result.

II. RELATED WORK

The algorithm we used for numerical prediction is based on model fusion. For a given numerical dataset of meteorology, MFVP firstly perform feature augmentation and feature selection on it, then use model fusion method to build predictor for visibility prediction after feature engineering. Besides, joint training were also been used for some specific meteorological station.

A. NUMERICAL PREDICTION OF VISIBILITY

Machine learning methods have been applied to numerical visibility prediction in recent years and some of them achieved good results. Current methods which use the result of the prediction to forecasting the weather are mentioned in the following.

One of these used methods is the regression algorithm. For example, [2] proposes a visibility forecasting system using the support vector machine algorithm (SVM) for Beijing Airport Expressway; [3] uses the support vector machine algorithm (SVM) to achieve short-term (0-12 hours) temperature forecasting, which is similar to visibility forecast, and compares the SVM scheme with the artificial neural network scheme; In [4], a logistic regression algorithm was used to establish a visibility prediction system for the Vienna airport in the cold season. Another method like in [9], the author proposes several Bayesian-based probabilistic visibility forecasting schemes. Except for regression algorithms, the neural network [11] also been used for numerical prediction. [5] builds a visibility forecasting system for Urumqi International Airport with deep neural networks(DNNs). And in considering that low visibility forecasts are of more significance, [12] proposes the Risk Neural Network, which focuses more on forecasting of low visibility. [13] uses DNN to build visibility forecasting system for 39 terminal aerodrome forecast stations in the northwest United States and compared their methods with logistic regression and model output statistics (MOS) derived from the Aviation Model/Global Forecast System. Furthermore, in considering that visibility forecasts are largely time-series dependent, [14], [15] proposed a visibility prediction system using Long-short term memory algorithm(LSTM).Moreover, some methods using multiple models for prediction also appeared, for example, several artificial neural networks are used to fuse the information of data to build a day-ahead electricity price prediction system in [16]; Some researches are done in [17] to value the performance of seven model selection criteria based on linear regression models with unknown noise variance.

As mentioned above, most visibility prediction systems based on machine learning algorithm using only meteorological numerical data. However, there are many other factors related to visibility haven't been considered yet.

B. MULTIMODAL FUSION

Information about a phenomenon or a system of interest can be obtained from different types of instruments, measurement techniques and other sources. It is rare that a single modality can provides comprehensive information to predict the tendency of the development of natural processes and environments [18]–[20]. The available access to multiple data enable the foundations of modern data fusion and the presentation of the multimodal. [21] proposed a multimodal stacked DPN (MM-SDPN), which consists of two-stage SDPNs, to fuse and learn highly feature representation from MRI (Magnetic Resonance Imaging) and PET (Positron Emission Computed Tomography) for AD (Alzheimer's disease) diagnosis; [22] proposed to take multimodal data as input of encoder, and acquire an affinity matrix corresponding to the data points in the self-expressive layer, and use decoder to reconstruct the input data, the DSC (Distributed Control System) network use reconstruction and the input data as training data to fuse the spatial information; [23]

proposed hierarchical-like multimodal grounded compositional semantics learning to model the inherent correlations between two modalities of collaboratively grounded semantics. As the aforementioned work illustrated, excellent performance is achieved by fusing the multimodal information. Theoretically, as two major modalities, the information of natural element and human beings' activities should improve the performance of the model, because the reason of the air quality degradation contains many elements including natural element and human beings' social activities.

C. SATELLITE IMAGES

The multimodal information can be provided by the satellite images since many applications of satellite images illustrated that usage of satellite image can assist the prediction [24]. The introduction of satellite image fits the thought of multimodal fusion by which excellent results has achieved.

As an innovative contribution, the estimation based on satellite images provides another modal to the multimodal fusion to assist the prediction of the visibility. Satellite images are digital images which contain massive information, those images are taken by a satellite vehicle through some sensors which perceive the specific wavelength of the light [25]. The advanced features of satellite images allow timely delivery of high-quality image data, high agility, and high accuracy for demanding operations [26], [27].

With the development of the satellite sensors, the latest generation of meteorological satellites can provide more frequent images with improved spatial resolution. Due to the authenticity and operability of satellite images, the demand for high-resolution satellite images has steadily increased [28]. The general application of satellite images is detections of vegetation, climatic disaster and the texture of the earth's surface.

Satellite image can be divided into three categories by resolution: low resolution, medium resolution and high resolution. Low-resolution satellite image is mainly used for global detection of climatic disaster and region climatic characteristics: [29], [30] proposed to use the satellite image to detect the center of tropical cyclone and the distribution of rainfall cloud cluster; [31]–[33] proposed the algorithm to detect the smoke without the disturbance of the thin cloud.

The short-wave infrared band is used in the sensor of the medium resolution satellite image. The infrared band is sensitive to the target with high temperature. Thus, the medium resolution satellite image is mainly used to detect the fire point and volcano. [34], [35] proposed the algorithm to detect the fire point with the low false positive rate.

With higher accuracy, the high resolution satellite image contains more precise information to detect the target with smaller sizes, such as buildings and ships [36], [37].

By analyzing the data characteristics of the medium resolution satellite image of Landsat-8 satellite, the traditional detection operator is introduced to estimate the emission of factories as the second modal information.



FIGURE 1. The distribution of meteorological monitoring stations.

TABLE 1. The 19 meteorological factors produced by every station.

Feature	Physical Meanings
p_08	Pressure
cha_p08	Pressure swing
t_08	Temperature
cha_t08	Variable temperature
rh_08	Humidity
rh850_08	850hPa humidity
rh700_08	700hPa humidity
rh_14	14 o'clock humidity
ud_dw	High and low altitude humidity difference
u_08	U-direction wind speed
v_08	V-direction wind speed
u850_08	850 hPa U-direction wind speed
v850_08	850 hPa V direction wind speed
w_ch	Wind shear
t850-t_08	Temperature difference between high and low altitude
wrh	Wind shear
t-td_08	The temperature-dew point depression
t08-td14	Variation of Dew Point Curve of Early-noon
see	Temperature
	Visibility

III. APPROACH

The approach proposed in this paper has two sub-modules. The numerical module proposes strategies to augment the meteorological features and improve the performance of the model. The strategies including adding the history data, selecting the feature, to determine whether to jointly train the features based on the distribution of meteorological stations. The output of the numerical prediction inputs as the first modal information.

Moreover, the estimation of the factory emission is introduced as the second modal information to assist the prediction of visibility. Instead of using the original image, some

preprocessing operations are proposed to highlight the target regions. The images after preprocessing are taken as the input to detect the location of the factory emissions and estimate the emission to shape the second modal information.

By fusing the output of the numerical prediction and the estimation, the final output turns out to better fit the real observation. The details of the approach are presented in the following.

IV. VISIBILITY PREDICTION BASED ON NUMERICAL METEOROLOGICAL FEATURES

Currently, most visibility prediction system is based on numerical meteorological data, and the construction of the visibility forecasting system mainly relies on relevant meteorological knowledge and experiences. Thus, it may cause the ignorance of some key information. Machine learning is proposed to complete the backbone of the numerical prediction system, including feature screening and predictive model learning.

Considering that the current machine learning algorithm model is relatively complete, so our work in the numerical prediction part mainly focuses on feature engineering and model fusion. The predicted meteorological data generated by each station every day for the next day is studied to predict the visibility of the next day. In the feature engineering part, instead of using traditional methods based on meteorological knowledge and subjective experience to screen features, more reliable machine learning methods are introduced to select appropriate features. In the part of model construction, a single machine learning algorithm is no longer to meet the demand, the model fusion strategy is introduced to fuse the prediction results of multiple algorithms. The detail of the numerical prediction method will be presented in this chapter.

A. DATA ANALYSIS AND PROCESSING

In this part, the data used for visibility prediction are meteorological monitoring data from more than 100 meteorological stations in the Beijing-Tianjin-Hebei region, the distribution of meteorological monitoring stations is shown in Figure 1. The meteorological data spanning about 15 years, from 2002 to 2017, each meteorological monitoring station generates data everyday, and each of which contains 19 meteorological factors recorded in that day. The 19 meteorological factors are given in Table 1.

It's obvious that the visibility and meteorological characteristics are not only spatially related, but also largely correlated with time shifts. Namely, the visibility is not only related to the meteorological characteristics of the current time but also related to the meteorological information in a certain historical time. So we further added meteorological information the day before for every single day. Thus we will get $19 \times (1 + 1) = 38$ meteorological factors for each station every day.

It can be seen from Table 1 that there are many kinds of meteorological factors. Through analysis, the influences

TABLE 2. The 10 factors selected according to the relevance ranking.

Feature	Physical Meanings
see_his1	Visibility of the previous day
t-td_08_his1	Temperature dew point difference of the previous day
t-td_08	The temperature-dew point depression
v850_08_his1	850 hPa V direction wind speed the day before
rh850_08	850 hPa humidity
p_08_his1	Pressure of the previous day
v850_08	850 hPa V direction wind speed
t850-t_08	Temperature difference between high and low altitude
ud_dw	High and low altitude humidity difference
p_08	Pressure

of all meteorological factors aren't equal to each other, and the correlation between each factor and visibility is also different [?]. Therefore, we first sort the correlation between meteorological factors (including historical meteorological characteristics) and visibility. The results are shown in Figure 2. We selected the top ten factors for training according to the relevance ranking. The selected factors are shown in Table 2.

B. ALGORITHM FOR NUMERICAL PREDICTION

The algorithm we used for numerical prediction is based on model fusion. For a given numerical dataset of meteorology, our method firstly performs feature augmentation and feature selection on it, then use the model fusion method to build a predictor for visibility prediction after feature engineering. Besides, joint training was also used for some specific meteorological station.

1) MODEL FUSION

Most current numerical visibility prediction schemes use only a single algorithm to train the model. [10] Considering the performance of algorithms for different data may be different either, two boosting tree algorithms, including Xgboost and LightGBM, are used to train the model separately for each site and perform model fusion.

Xgboost is integrated by many CART regression trees. Compared with previous boosting tree algorithm, the introduction of quadratic optimization make it possible to custom loss function for XGBoost, and its supporting of parallel computing has greatly increase the training speed.

LightGBM made further improvement base on XGBoost, it can get better trees with smaller computation cost and can avoid overfitting by using histogram algorithm and leaf-wise tree growth algorithm with max depth limitation.

In the early state research we found that XGBoost and LightGBM has different performance in different meteorological station, so model fusion was then introduced to our numerical visibility prediction scheme.

2) JOINT TRAINING FOR MULTI-STATION

The low visibility situation requires more attention in general. However, it is found through data analysis that there are

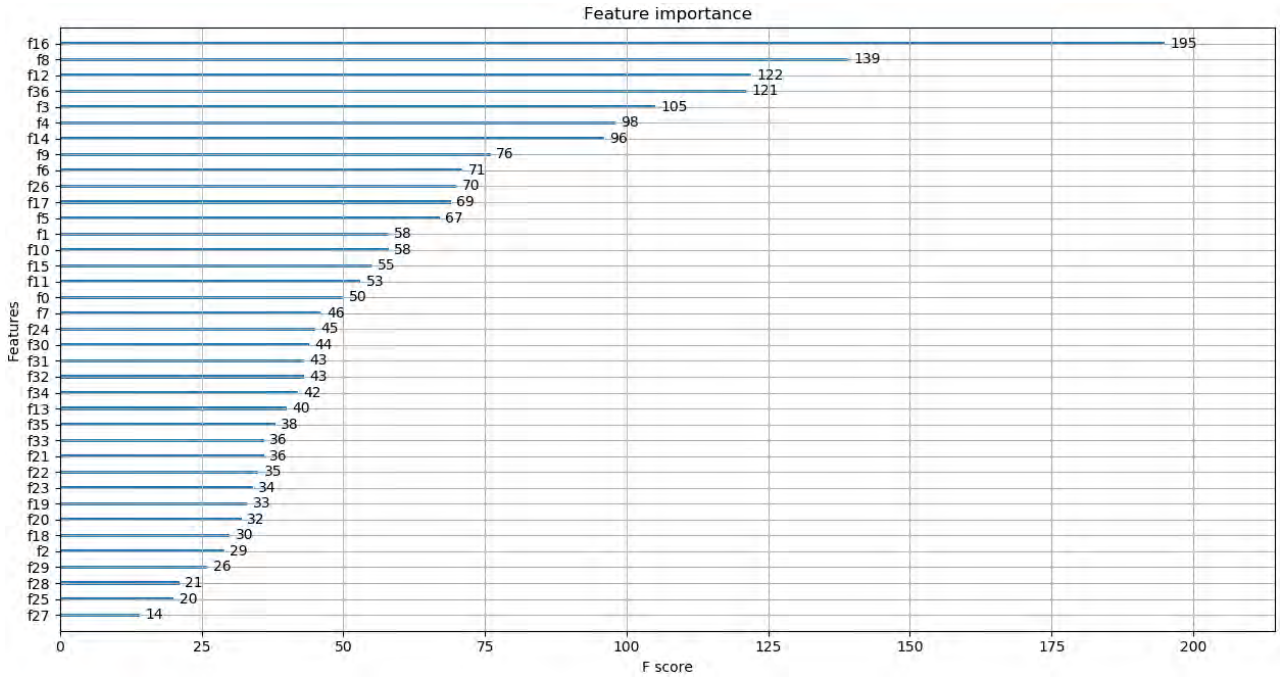


FIGURE 2. Ranking of correlations between different features and visibility.

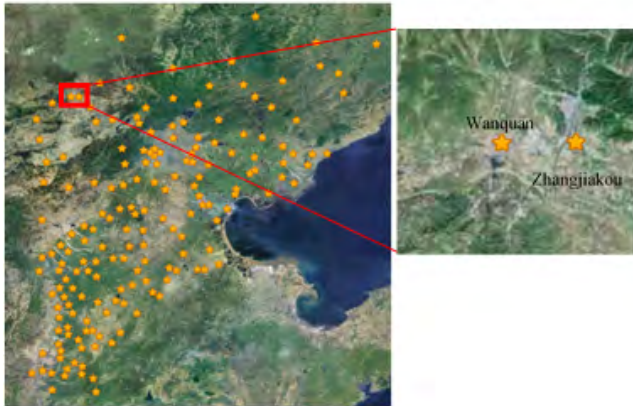


FIGURE 3. The location of Zhangjiakou station and Wanquan station.

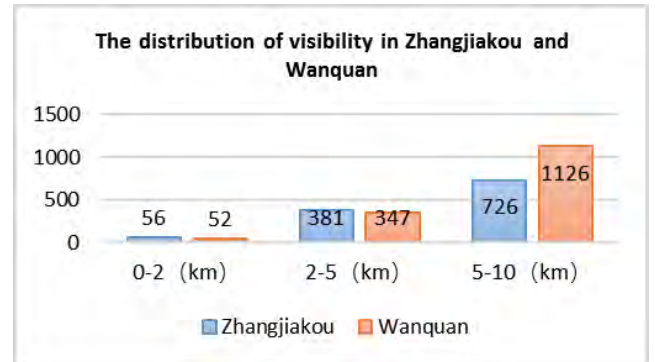


FIGURE 4. The distribution of visibility data in Zhangjiakou and Wanquan areas.

too few low-visibility data samples in some stations. And the absence of data samples leads to low accuracy for these stations. The distribution of the stations with low-visibility is dense, hence the joint train is proposed to improve the performance of models in these stations. Taking ZhangJiaKou station and WanQuan station as an example, the location of two stations is as shown in Figure 3 and the visibility distribution of these two stations can be seen in Figure 4. Visibility is divided into three levels ($0 < V \leq 2$, $2 < V \leq 5$, $5 < V \leq 10$, where V represent the visibility), and the vertical axis value indicates the number of samples in a certain range of visibility in an area. Models obtained by training two stations separately both get poor prediction accuracy, but by joint training, the models' performance can surpass the former models.

V. MULTIMODAL VISIBILITY PREDICTION

The emission of a large number of factories as a key factor together with natural factors affects urban visibility. Therefore, the estimation based on the satellite image is proposed to extract the second modal information, together with the first modal information to predict visibility.

A. DATASET USED

The Landsat-8 OLI/TIRS C1 Level-1 image date set used in our work is provided by Earth Explorer. The Landsat-8 launched by NASA carries two sensors: OLI (Operational land Imager, OLI) and TIRS (Thermal Infrared Sensor, TIRS). Landsat-8 maintains the same spatial resolution and spectral characteristics with Landsat 1-7. There are 11 bands. The spatial resolution of bands 1-7, 9-11 is 30 meters, and band 8 is 15 meters. Satellite Landsat-8 can

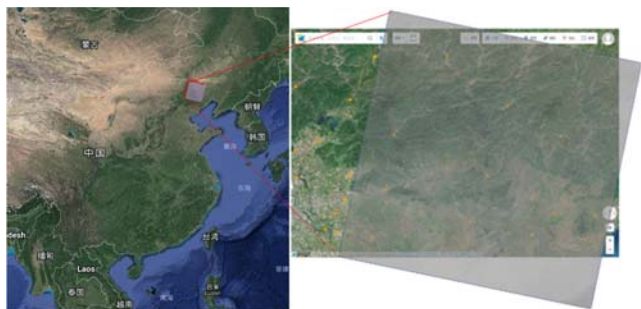


FIGURE 5. Distribution of meteorological stations in the selected region.

achieve global coverage every 16 days. The OLI Land Imager has 9 bands and the image width is 185×185 km. Compared with ETM sensor on Landsat-7, OLI Land Imager has made the following adjustments: (1). Band 5 has changed to $0.845\text{-}0.885 \mu\text{m}$, which excluding the influence of water vapor absorption on $0.825 \mu\text{m}$; (2). Band 8's panchromatic band is narrower, so the vegetation and no-vegetation areas can be better distinguished; (3). Two new bands are introduced. Band 1, namely blue band ($0.433\text{-}0.453 \mu\text{m}$) is mainly applied on coastal zone observation, and band 9, namely short-wave infrared band ($1.360\text{-}1.390 \mu\text{m}$) applied on cloud detection. The thermal infrared band carried on Landsat-8 is mainly used to collect heat loss in two thermal regions of the earth in order to understand the water consumption in the observed zones. The usage of Landsat-8 OLI/TIRS C1 Level-1 with the 10-TIRS band (30m) is surface temperature inversion, fire detection, soil moisture assessment, night imaging. The band used is $10.60 - 11.19 \mu\text{m}$. The 10-TIRS band is sensitive to the target with high temperature, such as factory emission points and high temperature emissions. The high temperature target presents as white spot on satellite images. Therefore, the Landsat 8 OLI/TIRS C1 Level satellite images with the 10-TIRS band are used to detect the location of factory emission point and estimate the emission range. The studied region is the northeast of Beijing. Located in the North China Plain, with dense distribution of factories, the region suffered haze for years. As shown in Figure 5, the meteorological stations in the studied region are densely distributed, and the estimated result can be reasonably distributed to the meteorological stations to fuse with the first modal data.

B. ESTIMATION OF FACTORY EMISSION

1) HIGHLIGHTING TARGET REGIONS

The size of satellite images from Landsat-8 OLI/TIRS C1 Level-1 with 10-TIRS is 7681×7801 . The characteristics of satellite images include the slope angle, the little proportion of factory emission points, the easily confused region of surrounding mountains. Considering the aforementioned characteristics, our work proposes a series of processes of "input - image preprocess - the detection of factory

emission point - estimate target - output the estimation" for satellite images.

After contour correction, the size of the satellite image is 6214×6551 . Because of the little proportion of the target, the grid segmentation is introduced in the preprocessing. The sub-image of the original infrared image is obtained by 10×10 grid segmentation. The satellite image to be detected is the sub-image.

The statistical histogram of the satellite sub-image is shown in Figure 6. As the image illustrated, it is hard to distinguish the target region from the surrounding by the pixel value. Therefore, the open operation is introduced to preprocess the input image. According to the traditional image segmentation method of preprocessing tumor image [38], [39], an open operation can highlighting the possible abnormal region in the image. As our experimental results presented, the open operation can well distinguish the background with the target region since those two regions both have high pixel value in the image. After the open operation, the binarization is implemented to segment the target from the background. The image after preprocessing has a more obvious target region and easier to be detected and estimated.

2) EMISSION DETECTION AND ESTIMATION

The output of the preprocessing is used as input in this part. Given the acquisition cycle of satellite images is 16 days, and without the label, the amount of data can be processed is small. Thus, we propose an algorithm based on traditional SURF detector for target detection and estimation for factory emissions. Since SURF (Speeded Up Robust Features, SURF) constructs a pyramid of scale space in the process, it can adapt to the scale translation of the target region well. Compared to SIFT, SURF's descriptor can well meet the demand of our work since it contains the information of the target's diameter and location. Thus, the SURF is introduced to detect and estimate the emission. The output of SURF detection has a large number of overlapping regions. The de-duplication is implemented to obtain the region with the largest coverage. The data analysis indicates that the difference between the target region and the surrounding can be determined by the pixel value and texture. With the fixed range of threshold, part of the candidate region will be removed. For the left candidate target region, the Sobel is used to detect the texture of the candidate region. The candidate region with a smooth texture is the target region as shown in Figure 6. Since the 10×10 grid segmentation is applied in the preprocessing, the emissions can be assigned into the meteorological station in the corresponding grids as the second modal information.

VI. MULTIMODAL FUSION

The multimodal information is essential to introduce for analyzing the complicated phenomenon like the atmospheric visibility. The key factors contributed to the low atmospheric visibility are diverse. As the innovative contribution, in this chapter, the multimodal fusion is presented by combining the

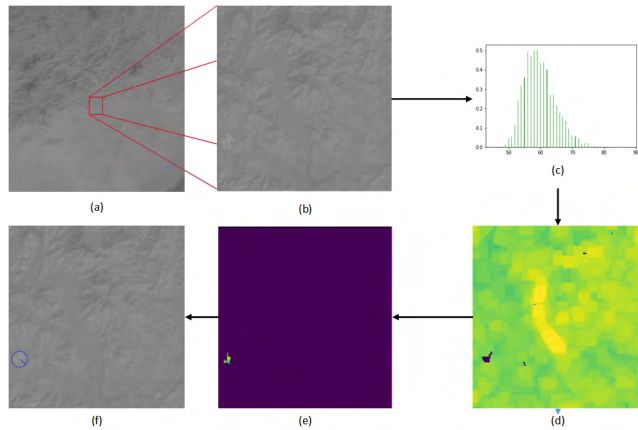


FIGURE 6. The result of image processing. (a) Overview of the study region. (b) The study region after grding. (c) The histogram of study region. (d) Diagram of corrosion operation. (e) Diagram of binary operation. (f) Final result of image processing.

information of the numerical prediction and the estimation of the factory emissions to further improve the performance of the proposed system. The multimodal fusion proposed can be illustrated as the following equation:

$$y = k_1x_1^2 + k_2\bar{x}_2. \tag{1}$$

In the equation, the y is the final result of the prediction of the visibility., the x_1 presents the output of the numerical prediction of visibility and the x_2 presents the estimation of the factory emission. Given the fact that the emission at some region will be low even with the low visibility. In this region, the estimation is supposed to have a slight influence. With such inference, the \bar{x}_2 presents the normalized estimation and the influence of numerical prediction is relatively augmented by the operation. To further augment the influence of numerical prediction, the output of numerical prediction is squared in the equation, as x_1^2 . With such operation, the estimation of the factory emission work as a residual factor to assist the prediction of visibility. By application of linear regression to the output of the numerical prediction and the estimation of factory emissions, as the Figure 7 presented result of the prediction fuses the two modal information to obtain more accurate performance.

VII. EXPERIMENTS

A. EVALUATION CRITERIA

The level of visibility is divided into 6 levels from 0 to 5 based on the meteorological industry standard of the People’s Republic of China Considering that the high visibility has little impact on our life, high attention is placed on the forecast of weather with lower visibility. The low visibility weather is re-divided into 4 levels, as Table 3 illustrated. According to the Visibility forecasting grade and serving [40] issued by the National Weather Service, this paper uses the TS score to evaluate the forecast results. The TS score is computed as:

$$TS = \frac{n_a}{n_a + n_b + n_c}. \tag{2}$$

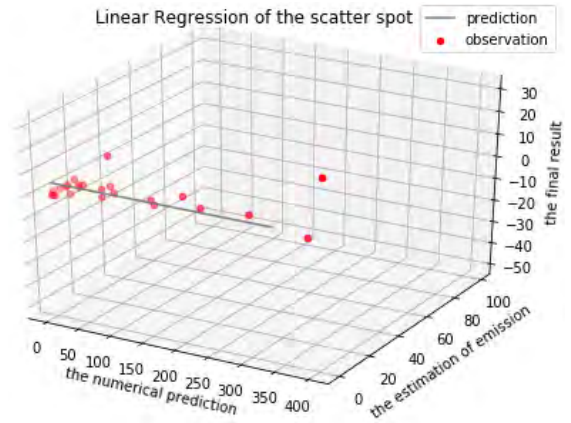


FIGURE 7. The result of linear regression for the proposed method.

TABLE 3. The rating of weather visibility observations.

Observations	Rating
$V > 10$	The visibility is in excellent situation.
$5 < V \leq 10$	The visibility is in relatively good situation.
$2 < V \leq 5$	The visibility is in relatively bad situation.
$0 < V \leq 2$	The visibility is in poor situation.

where n_a is the number of correct predictions (When the forecasting visibility is ranked the same as the actual visibility, the forecast is correct.), n_b is the number of empty reports (For a certain grading, when the forecast result is within the grading while the actual value isn’t, then it is recorded as an empty report.), n_c is the number of missing reports(For a certain grading, when the forecast result is not within the grading while the actual value is, then it is recorded as a missing report.).

In order to further evaluate the accuracy of the predicted results, we also take the RMSE (Root Mean Square Error) into account:

$$E_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - R_i)^2}. \tag{3}$$

where N is the total number of samples that we predict, P_i is the predicted result for i -th sample, R_i is the actual value for i -th sample.

B. EVALUATION ANALYSIS

1) NUMERICAL DATA BASED PREDICTION

Firstly, we tested the accuracy of our method (model fusion of XGBoost and LightGBM) for numerical data based prediction and compared it with several existing numerical prediction methods, including XGBoost, LightGBM, SVM and ANN method. The meteorological data from 2002 to 2016 were used for training and meteorological data of 2017 were used for testing. The experimental area is as mentioned before. After feature engineering and feature screening, we retained 10 meteorological factors. And we use the same data processing method for each algorithm for

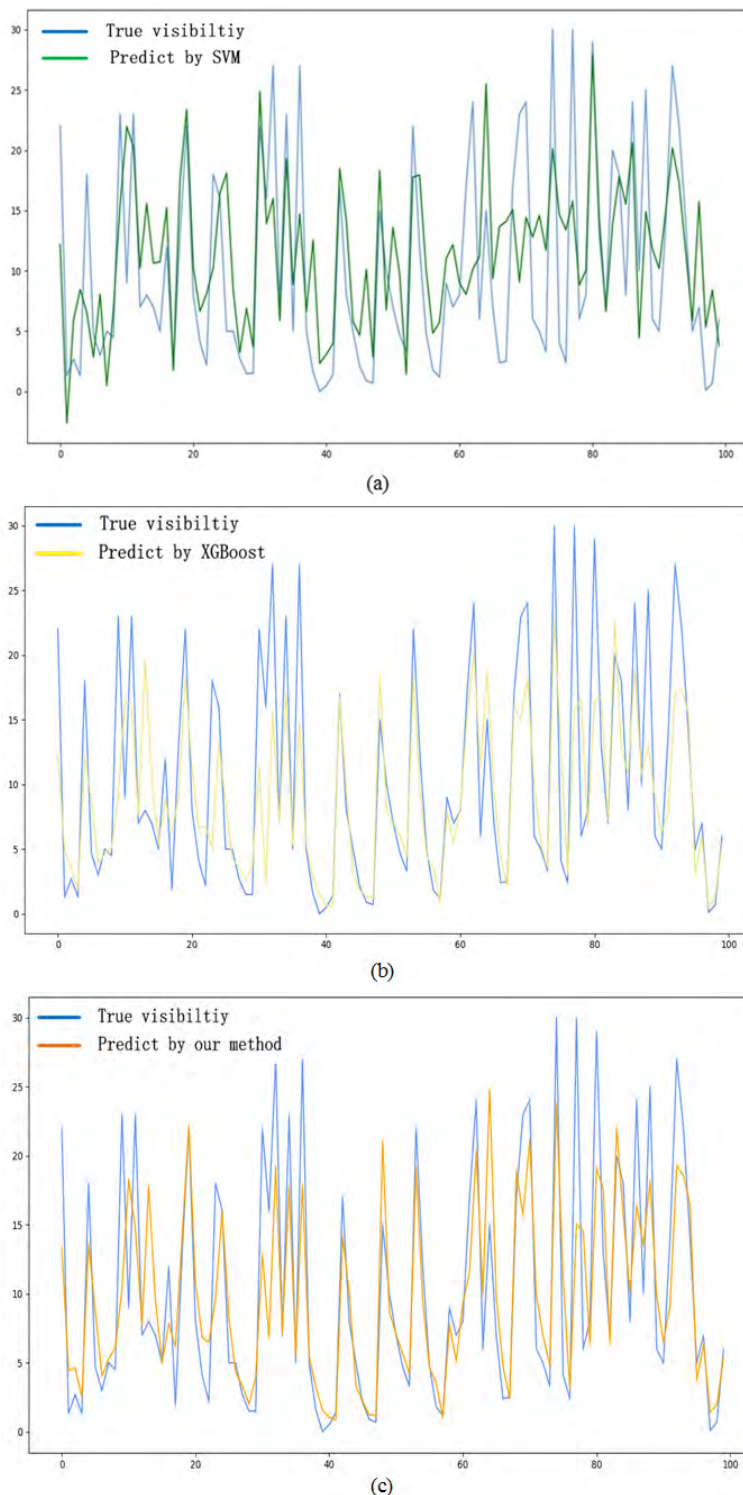


FIGURE 8. The prediction results compared with truth. We took the last 100 days’ data from meteorological observation station in Beijing for testing, figure (a), (b), (c) compares the visibility predicted by SVM, XGBoost and our methods with truth value respectively.

fairness. The scoring criteria are the TS scores and RMSE described in section A. The result is shown in Table 4, and taking the meteorological observation station in Beijing as an example, we plot the true visibility value and prediction results generated by several methods in Figure 8.

2) EMISSION DETECTION AND ESTIMATION BASED ON SATELLITE IMAGERY

Since the amount of Landsat 8 OLI/TIRS C1 Level-1 satellite images with 10-TIRS is small, the traditional detection is more suitable for the satellite images. The different detection

TABLE 4. The rating of weather visibility observations experimental results of each algorithm for numerical data based visibility prediction. we focus on the results in visibility range of 0-10 km.

Methods	0-2 (km)	2-5 (km)	5-10 (km)	RMSE
XGBoost method	0.29	0.15	0.16	8.09
LightGBM method	0.16	0.13	0.17	7.96
SVM method	0.17	0.12	0.16	8.53
ANN method	0.085	0.10	0.15	8.01
MFVP	0.29	0.16	0.18	7.38

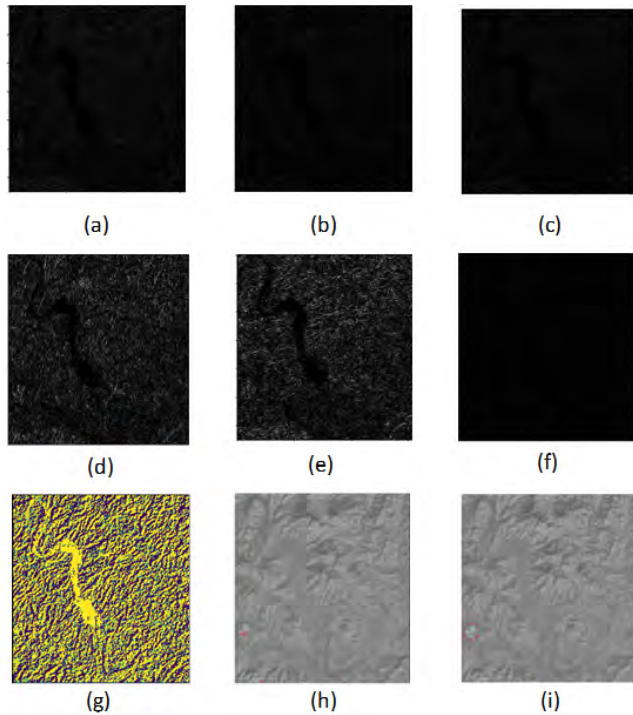


FIGURE 9. Fig(a)-(c) is the result of SobelCombine, SobelX, SobelY detector; Fig(d)-(e) is the result of ScharX and ScharY detector; Fig(f) is the result of Laplace detector; Fig(g) is the result of LBP detector; Fig(h) is the result of SIFT detector; Fig(i) is the result of SURF detector.

algorithms are discussed in Figure 9. Through the analysis of the images above, the conclusion can be made as following: the Sobel detector can detect the stripe-like geographic features of Landsat 8 OLI/TIRS C1 Level-1 satellite images with 10-TIRS band; Schar detector is more sensitive to geographic features, but cannot tell the difference between the target and the surrounding; Laplace detector is not sensitive to the features in the satellite images; LBP (Local Binary Pattern, LBP) detector can highlight the texture of background without telling the difference between the target with the surroundings; SIFT (Scale-invariant feature transform, SIFT) detector can accurately detect the location of the factory emission points without the effectively estimable emission range of the factory emission; SURF detector can successfully detect the factory emission point with accurate location and the estimable emissions range. Through the detection and the estimation, the result emission range is presented by the number in the 10×10 grids. With such design, the result can

TABLE 5. Results of comparative experiments.

Methods	RMSE
XGBoost	8.63
LightGBM	8.47
ANN	8.52
Our Methods	8.35
MFVP	6.71

be assigned to the corresponding meteorological stations to assist the prediction of the visibility.

3) EMISSION DETECTION AND ESTIMATION BASED ON SATELLITE IMAGERY

Multimodal fusing was proposed after getting results from numerical prediction and emission detection. We further conducted comparative experiments to test the effects of multimodal fusion.

a: DATA

The numerical data we used in this part was generated by 22 meteorological observation stations of the target region mentioned before in 2018, they are structurally the same as the data we used in the previous section. Numerical data before 2018 was used for model training and data of 2018 was used for the test. The satellite images used for emission estimation also come from Landsat 8 OLI/TIRS C1 Level-1 with 10-TIRS in 2018.

b: EXPERIMENTAL DETAILS

By fusing preliminary prediction results and emission estimation results we make regression predictions as proposed. To verify the effectiveness of our method, we compared MFVP (Multimodal Fusion for Visibility Prediction) with several numerical-data-only methods. Using root mean square error(RMSE) of prediction results for evaluating, the results are as shown in Table 5. According to the results shows in Table 5, it can be seen that the RMSE of prediction results generated by MFVP is significantly better than RMSE of other numerical-data-only methods. That is to say, by fusing multimodal information, the visibility prediction system can be significantly improved.

VIII. CONCLUSION

A multimodal fusion method for visibility prediction was proposed in this paper. It can be concluded from our experiments that by fusing the numerical prediction with estimation of the factory emission based on satellite images, the accuracy of visibility prediction was significantly improved. As the further inference, in addition to the emission of factories, other factors that have an impact on visibility can also be introduced by multimodal fusion method to improve the prediction system. Furthermore, problems that have something in common with visibility prediction, such as weather forecasting and PM 2.5 prediction, can be solved better by introducing multimodal fusion. There are also some disadvantages in our MFVP (Multimodal Fusion for Visibility Prediction). For

example, the emission detection requires the satellite images have high resolution, however, it's not easy to get high resolution satellite images for a certain region, which may cause lacking data for prediction using our system. There are still many ways to solve this problem, such as introducing other modality to replace satellite images. Furthermore, the factor that taking into account was only the emission of factories in this paper, more works can be done by taking more factors into consideration in the future to further improve our visibility prediction system [41].

REFERENCES

- [1] Y. He, X. Sun, L. Gao, and B. Zhang, "Ship detection without sea-land segmentation for large-scale high-resolution optical satellite images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 717–720.
- [2] Z.-W. Wang, C.-L. Zhang, C. Su, and C.-L. Cheng, "On modeling of atmospheric visibility classification forecast with nonlinear support vector machine," in *Proc. 5th Int. Conf. Natural Comput.*, Aug. 2009, pp. 240–244.
- [3] R. F. Chevalier, G. Hoogenboom, R. W. McClendon, and J. A. Paz, "Support vector regression with reduced training sets for air temperature prediction: A comparison with artificial neural networks," *Neural Comput. Appl.*, vol. 20, no. 1, pp. 151–159, 2011.
- [4] P. Kneringer, S. J. Dietz, G. J. Mayr, and A. Zeileis, "Probabilistic nowcasting of low-visibility procedure states at vienna international airport during cold season," *Pure Appl. Geophys.*, vol. 176, no. 5, pp. 2165–2177, 2018.
- [5] Z. Lei, G. Zhu, H. Lei, and W. Nan, "The application of deep learning in airport visibility forecast," *Atmos. Climate Sci.*, vol. 7, no. 3, pp. 314–322, 2017.
- [6] X. Rao, L. Qiao, and X. Wang, "Visibility climate characteristics and influencing factor analysis during Beijing national day," (in Chinese), *Weather Forecast Rev.*, vol. 2, no. 1, pp. 10–16, 2010.
- [7] X. Yu, J. Ma, J. An, L. Yuan, B. Zhu, D. Liu, J. Wang, Y. Yang, and H. Cui, "Impacts of meteorological condition and aerosol chemical compositions on visibility impairment in Nanjing, China," *J. Cleaner Prod.*, vol. 131, pp. 112–120, Sep. 2016.
- [8] Y. Bai, H. Qi, L. Liu, C. Chen, C. Lin, and W. Li, "Study on the nonlinear relationship among the visibility, PM_{2.5} concentration and relative humidity in Wuhan and the visibility prediction," (in Chinese), *Acta Meteorologica Sinica*, vol. 74, no. 2, pp. 189–199, 2016.
- [9] R. M. Chmielecki and A. E. Raftery, "Probabilistic visibility forecasting using Bayesian model averaging," *Monthly Weather Rev.*, vol. 139, no. 5, pp. 1626–1636, 2011.
- [10] D. Bari, "Visibility prediction based on kilometeric nwp model outputs using machine-learning regression," in *Proc. IEEE 14th Int. Conf. e-Sci.*, Nov. 2018, p. 278.
- [11] Z. Ma, J.-H. Xue, A. Leijon, Z.-H. Tan, Z. Yang, and J. Guo, "Decorrelation of neutral vector variables: Theory and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 129–143, Jan. 2018.
- [12] W. Kai, Z. Hong, A. Liu, and Z. Bai, "The risk neural network based visibility forecast," in *Proc. 5th Int. Conf. Natural Comput.*, Aug. 2009, pp. 338–341.
- [13] C. Marzban, S. Leyton, and B. Colman, "Ceiling and visibility forecasts via neural networks," *Weather Forecasting*, vol. 22, no. 3, pp. 466–479, 2007.
- [14] K. Abhishek, M. Singh, S. Ghosh, and A. Anand, "Weather forecasting model using artificial neural network," *Procedia Technol.*, vol. 4, pp. 311–318, Jan. 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S221201731200326X>
- [15] Y.-T. Tsai, Y.-R. Zeng, and Y.-S. Chang, "Air pollution forecasting using RNN with LSTM," in *Proc. IEEE 16th Int. Conf. Dependable, Autonomous Secure Comput., 16th Int. Conf. Pervasive Intell. Comput., 4th Int. Conf. Big Data Intell. Comput. Cyber Sci. Technol. Congr.*, Aug. 2018, pp. 1074–1079.
- [16] H. Tian, K. Li, and B. Meng, "A novel prediction modeling scheme based on multiple information fusion for day-ahead electricity price," in *Proc. Chin. Control Decis. Conf. (CCDC)*, May 2011, pp. 1801–1805.
- [17] H. Chen and S. Huang, "A comparative study on model selection and multiple model fusion," in *Proc. 7th Int. Conf. Inf. Fusion*, Jul. 2005, p. 7.
- [18] R. Ji, F. Chen, L. Cao, and Y. Gao, "Cross-modality microblog sentiment prediction via Bi-layer multimodal hypergraph learning," *IEEE Trans. Multimedia*, vol. 21, no. 4, pp. 1062–1075, Apr. 2019.
- [19] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.
- [20] T. Chen, S. Wang, and S. Chen, "Deep multimodal network for multi-label classification," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 955–960.
- [21] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 173–183, Jan. 2018.
- [22] M. Abavisani and V. M. Patel, "Deep multimodal subspace clustering networks," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1601–1614, Dec. 2018.
- [23] Y. Zhuang, J. Song, F. Wu, X. Li, Z. Zhang, and Y. Rui, "Multimodal deep embedding via hierarchical grounded compositional semantics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 1, pp. 76–89, Jan. 2018.
- [24] X. Zhao, L. Gao, Z. Chen, B. Zhang, W. Liao, and X. Yang, "An entropy and MRF model-based CNN for large-scale landsat image classification," *IEEE Geosci. Remote Sens. Lett.*, to be published.
- [25] A. Sağlam and N. A. Baykan, "Roof detection on satellite images," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, May 2017, pp. 1–4.
- [26] E.-E. Kim, H. Lee, and E. D. Kim, "Sub-half meter imaging satellite, SpaceEye-X," in *Proc. 8th Int. Conf. Recent Adv. Space Technol. (RAST)*, Jun. 2017, pp. 125–127.
- [27] Z. Ma, Y. Lai, W. B. Kleijn, Y.-Z. Song, L. Wang, and J. Guo, "Variational Bayesian learning for Dirichlet process mixture of inverted Dirichlet distributions in non-gaussian image feature modeling," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 449–463, Feb. 2019.
- [28] I. Demir, K. Koperski, D. Lindenbaum, G. Pang, J. Huang, S. Basu, F. Hughes, D. Tuia, and R. Raskar, "DeepGlobe 2018: A challenge to parse the earth through satellite images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 172–181.
- [29] L. Jia, C. Liu, B. Wang, and D. Qin, "A novel algorithm for detecting center of tropical cyclone in satellite infrared images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 917–920.
- [30] M. A. Tebbi and B. Haddad, "Improving infrared MSG satellite images classification, application on rainy areas detection," in *Proc. 8th Int. Conf. Model. Identificat. Control (ICMIC)*, Nov. 2017, pp. 747–750.
- [31] M. Sheng, K. Hu, G. Hao, and X. Wang, "Small fire smoke region location and recognition in satellite image," in *Proc. 9th Int. Congr. Image Signal Process., Biomed. Eng. Inform.*, Oct. 2016, pp. 714–718.
- [32] S. Miao, H. Lin, H. Gao, and L. Dong, "Strip smoke and cloud recognition in satellite image," in *Proc. 9th Int. Congr. Image Signal Process., Biomed. Eng. Inform.*, Oct. 2016, pp. 303–307.
- [33] B. P. Shukla and P. K. Pal, "Automatic smoke detection using satellite imagery: Preparatory to smoke detection from INSAT-3D," *Int. J. Remote Sens.*, vol. 30, no. 1, pp. 9–22, 2009.
- [34] H. E. Quanjun and L. I. U. Cheng, "Improved algorithm of self-adaptive fire detection for MODIS data," *J. Remote Sens.*, vol. 12, no. 3, pp. 448–453, 2008.
- [35] M. J. Pavolomis, J. Sieglaff, and J. Cintineo, "Automated detection of explosive volcanic eruptions using satellite-derived cloud vertical growth rates," *Earth Space Sci.*, vol. 5, no. 12, pp. 903–928, 2018.
- [36] A. Zakharov, A. Tuzhilkin, and A. Zhiznyakov, "Automatic building detection from satellite images using spectral graph theory," in *Proc. Int. Conf. Mech. Eng., Automat. Control Syst.*, Dec. 2015, pp. 1–5.
- [37] J. L. Wang and X. L. Liu, "Evaluation of a new visiometer for automated visibility observation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 3481–3484.
- [38] L. N. Chen and L. Yang, "New method on liver tumor CT image segmentation," (in Chinese), *Comput. Eng. Appl.*, vol. 45, no. 34, pp. 222–224, 2009.
- [39] C. Feng, H. Chunbao, and Z. Lihui, "Research on cell image preprocessing algorithm based on tissue slices," (in Chinese), *Electron. World*, vol. 20, pp. 142–142, 2014.
- [40] China Meteorological Administration, *QXT 114-2010 Meteorological Industry Standards of the People's Republic of China*. Beijing, China: Meteorological Publishing House, 2010, no. 46.

- [41] Z. Ma, J. Xie, Y. Lai, J. Taghia, J. Xue, and J. Guo, "Insights into multiple/single lower bound approximation for extended variational inference in non-Gaussian structured data modeling," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.



CHUANG ZHANG (M'18) was born in Changchun City, China, in 1975. He received the B.S. degree in mechanical engineering from the Jilin University of Technology, Changchun, China, in 1998, and the M.S. degree in mechanical engineering from Jilin University, in 2001, the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications (BUPT), in 2004. Since 2006, he has been an Associate Professor in the Lab of Pattern Recognition and Intelligence System (PRIS), School of Information and Telecommunications Engineering, BUPT. His research interest includes computer vision and pattern recognition, satellite image parsing, and human gesture recognition. He has published over 50 technical papers in international journals and conferences.



MING WU was born in Shanghai, China, in 1977. She received the M.S. and Ph.D. degrees in Information and Communication Engineering from the Beijing University of Posts and Communications, Beijing, China, in 2003 and 2012, respectively. She is currently an Associate Professor with the School of Information and Communication Engineering, Beijing University of Posts and Communications, China. Her primary research interests include computer vision, pattern recognition, deep learning and satellite image parsing, and human gesture recognition. She has published over 20 technical papers in international journals and conferences.



JINYU CHEN was born in Fujian, China, in 1996. She has studied the Internet of Thing at Tianjin University. She obtained the B.S. degree from Tianjin University. She is currently pursuing the master's degree in electronic information engineering with Beijing University of Posts and Telecommunications. She is a member of the Pattern Recognition and Intelligent System Lab. Her research interests include detection based on the satellite image and data augmentation.



KAIYAN CHEN was born in Ningde, China, in 1996. She received the B.S. degree in information and communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2018. She is currently pursuing the master's degree in information and communication engineering with the Beijing University of Posts and Telecommunications. Her current research interests include computer vision, pattern recognition, and deep learning. She is now a member of the Lab of Pattern Recognition and Intelligence System (PRIS), School of Information and Telecommunications Engineering, BUPT.



CHI ZHANG received the bachelor's degree in Peking University in China, and the master's degree in climate economics in Stockholm University, and the Ph.D. degree in energy processes from Royal Institute of Technology - KTH, Sweden. She is currently working as a Postdoctoral Fellow with the topics of energy economics and climate policy, environmental economics and emission trading, innovation and sustainable development, and healthy economics in KTH.



CHAO XIE was born on Dec 30, 1987 in Beijing, China. He completed the master's degree in environmental science and engineering from the Beijing University of Technology. He works as a synoptic with National Meteorological Center, Beijing. His main research interests include the fog-haze-visibility forecast and model output statistics.



BIN HUANG received the bachelor's degree in China Ocean University, and the master's degree in the Nanjing University of Information Engineering. She is currently working with the National Meteorological Center of China Meteorological Administration, engaged in marine meteorological forecasting technology research work.



ZICHEN HE is currently pursuing the bachelor's degree in statistics and computer Science with McGill University, Montréal, Canada. She works with Professor Nikolas Provatas as a teaching assistant in McGill University. Also, she did some projects about designing AI for chess game and simulating automation system. Her current research interests include using data from social platforms, human communities, political parties, and ancient civilizations to measure and predict large-scale human behavior.

...