A Multi-source Remote Sensing Identification Method for Coal Fire Zones Based on a Three-layer Ensemble Learning Model

Yu Chen, Kaiwen Ding, Zhihui Suo, Kun Tan, Jun Wei, Fei Cao, Huahai Sun, Mengmeng Bu

[1] ***Abstract*—Spontaneous combustion of coal seams poses significant threats to ecology and human health, making it a global concern. Precise location identification of coal fire burning is vital for disaster management. Traditional methods of identifying fire zones focus on specific indicators, such as temperature anomalies. However, solely relying on one indicator can easily lead to "heterogeneous homogeneity", and using multiple indicators complicates threshold settings. To address this, our study presents an enhanced ensemble learning model named AdaBoost_RF_StBP for coal fire zone identification. By integrating multi-modal remote sensing data, we extracted diverse abnormal features of fire zones. Using a standard deviation ellipse, we categorized feature samples into three fire zone types, which were then analyzed using the proposed model. The combined approach of base and meta learners was applied in coal fire zones located in Jiangjun Gobi, Jiangjun Temple, Sandaoba, and Beishan coal mines. Results show that the multi-modal remote sensing data coupled by our model can effectively addresses the "heterogeneous homogeneity" issue. The accuracy of the model in terms of accuracy, precision, recall rate, F1_Score, etc., has been improved from 0.56-0.87 of the initial base-learners to 0.8-0.96 of the meta-learners. Additionally, the recognition areas of AdaBoost_RF_StBP in Jiangjun Gobi, Jiangjun Temple and Sandaoba fire areas account for 122.7%, 109.8% and 115.8% of the actual fire areas, respectively. This underscores the promising potential of the AdaBoost_RF_StBP model for fire zone identification. This model offers valuable insights and guidance for large-scale fire zone monitoring and identification, proving beneficial for precise and efficient fire control.**

***Index Terms*—Coal fire identification, Multi-source remote sensing, Machine learning, Ensemble learning**

I. Introduction

COALFIELD fire zones, which are expansive areas resulting from the sustained combustion of surface coal outcrops, can extend from shallow to deep, and their formation and development are influenced by the combination of natural conditions and human activities [1], [2], [3]. As an economic resource, coal is being extensively exploited. Unregulated mining exacerbates coal oxidation and spontaneous combustion, leading to the expansion of coalfield fire zones. This is particularly serious in countries and regions rich in coal resources such as the United States, Russia, Australia, China, India and South Africa [4], [5], [6]. Coal fires not only deplete coal resources but also emit harmful substances, including coke tar, sulfur, and carbon. They also release gases that have an incalculable impact on the surrounding ecosystem, water resources, soil, and atmospheric environment [7], [8], [9]. Therefore, how to quickly and accurately identify and delineate the location of coal fire zones has become the focus over the years.

The intense underground combustion of coal alters the physical and chemical landscapes in fire areas. Therefore, four major types of field detection methods, including geophysical, geochemical, thermal, and drilling, represented by surface temperature measurement [10], natural potential [11], transient electromagnetics [12], isotope measurement of radon and sulfur [13], [14], and binary tracer gas [15] have been proposed successively. With the significant improvement of the spatial-temporal resolution and performance of satellite-borne sensors, the use of remote sensing data to characterize coal fire combustion and subsequent identification has been widely studied. This mainly includes thermal infrared monitoring [10], [16], [17], subsidence monitoring [18], [19], [20] and multi-source remote sensing cooperative monitoring [21], [22], [23], [24]. Tien et al. [25] used Lansat-8 TIRS data to invert surface temperature based on radiation transfer equations and combined hotspot analysis to delineate the Khanh Hoa coal fire zone in northeastern Vietnam. Jiang et al. [26] investigated the feasibility and potential of detecting the land subsidence accompanying coal fires by means of satellite InSAR observations, which the interferometric results agree well with GPS observations and coal fire data obtained by field investigation. On this basis, Jameela et al. [27] used New Small Baseline Subset InSAR (N-SBAS) technology to accurately estimate vertical subsidence in Kusunda Coal Mine in India. While reducing the occurrence of temporal-spatial coherence loss in dynamic areas such as coalfield fire zones, it has a good correspondence with Landsat-8 satellite temperature anomaly maps and field evidence. Wang et al. [28] used strong-weak joint constraints on multi-source remote sensing data that reflected different phenomena of coal fire combustion. They constructed band-pass filters for fire zone identification, and achieved a 91% accuracy rate in identifying the Fukang fire zone in Xinjiang. Yu et al. [29] integrated multi-temporal thermal infrared and radar data to determine temperature anomaly frequency thresholds, and developed a Coal Fire Ratio Index (CRI) for quantitatively

This work was supported by the National Natural Science Foundation of China (Grant Numbers: 42171312), the Basic Research Program of Xuzhou (Grant Numbers: KC23049), and the Open Fund of Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University (Grant Numbers: KLGIS2023A02). *(Corresponding author: Yu Chen)*

Yu Chen, Kaiwen Ding, Zhihui Suo are with the Key Laboratory for Land Environment and Disaster Monitoring, Ministry of Natural Resources, China University of Mining and Technology, Xuzhou 221116, China (e-mail: chenyu@cumt.edu.cn; kevintyn@163.com; SuoZH@cumt.edu.cn).

Kun Tan is with the Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China (e-mail: tankuncu@gmail.com).

Jun Wei, Fei Cao, Huahai Sun are with the Xinjiang Uygur Autonomous Region Mine Safety Service and Guarantee Center, Urumqi 830017, China (e-mail: weijunwxx@163.com; figo6507804@163.com; 297228108@qq.com).

Mengmeng Bu is with the Fifth Geological Brigade of the Bureau of Geology and Mining, Xuzhou 221004, China (e-mail: bumengcumt@gmail.com).

assessing the severity and changes of coal fires. Wang et al [30] proposed a Spatio-Temporal Temperature-Based Thresholding algorithm to enhance thermal anomaly detection capability and introduced Polarimetric Persistent Scatterer Interferometry to enhance the interference effect in fire areas. Chen et al [31] proposed a Temporal Temperature Anomaly Extraction algorithm based on Adaptive Windows (TTAE-AW) and integrated it with DS-InSAR to propose a Double Threshold Two-stage Filter (DTTF). In practice, the temperature anomalies extracted by TTAE-AW are more concentrated in the actual fire areas, and the DTTF has a better identification performance compared to various traditional threshold methods.

Each of the aforementioned coal fire monitoring methods possesses unique advantages and has shown practical results. However, there are still many shortcomings. While field detection methods offer reliable measurement results, they tend to be inefficient, limited in scope, and offer lower spatial resolution. The limitation of single satellite-borne thermal infrared detection is the ambiguity of defining temperature thresholds for distinguishing background pixels from coal fire pixels. Due to the influence of surface cracks, the detected thermal anomalies might not directly indicate coal fires in the vertical direction [16]. The limitation of using deformation anomaly to detect coal fire zones is that it is easily affected by disturbances induced by other deformation sources, e.g. mining subsidence [32]. Multi-source remote sensing information filtering can take into account the above-mentioned situations to a certain extent. However, due to the complex geological conditions in the fire area, the apparent abnormal position does not overlap vertically with the actual fire area. The asynchronous multi-source information from burning zones and the concealed nature of deep surface fires result in the burning and developmental state being invisible. The intermediate parameters of the filter itself need to be adjusted for different periods and different fire areas, and the threshold setting is highly subjective. Considering all these factors, the filtering method sometimes cannot effectively eliminate interference information. In addition, due to the lack of clear physical connections between multi-source data sets, such methods also do not address issues like decreased sensitivity in certain parameters. This leads to a reduction in the generalizability and continuous monitoring capability of the method across different regions.

Where statistical models struggle with multidimensional data sets, machine learning can offer solutions. Its goal is to use actual data sets to construct a prediction model with strong generalization ability, so as to give as accurate an estimate as possible for unknown objects [33], [34]. However, in practice, it is often difficult for a single learning model to find a suitable model in the assumed space. Ensemble learning, which employs multiple learners to tackle the same problem, boasts enhanced model generalization capabilities. Its idea can be roughly divided into heterogeneous ensemble that applies different learning algorithms to the same data sets and homogeneous ensemble that applies the same learning algorithm to different data sets [35], [36]. At present, most applications of ensemble learning employ a single ensemble learning approach, lacking a comparison between different ensemble methods. Moreover, very limited research on coal fire zone identification using ensemble learning has been conducted.

To more effectively discern the potential relationships among multi-source remote sensing data in fire zones, we integrated thermal infrared, multispectral, SAR, and night light data to extract multi-modal remote sensing information under strong constraints. We combined Boosting, Bagging, and Stacking ensemble ideas to construct an improved three-layer ensemble learning model, named AdaBoost_RF_StBP, for coal fire zone identification. This approach leverages the advantages of different ensemble methods and improves the robustness, generalizability, and accuracy of the model and identification results. We verified the model's accuracy and identification capabilities using three test areas in Jiangjun Gobi, Jiangjun Temple (including Beishan coal mines) in Qitai, and Sandaoba in Miquan County, Xinjiang, respectively. This provides a novel approach for coal fire zone identification.

II. Methodology

In response to the existing problems outlined earlier regarding fire zone identification, this study proposes a multi-source remote sensing identification method for coal fire zones based on a three-layer ensemble learning model. Firstly, we used multi-modal remote sensing data including thermal infrared, multispectral, SAR and night light data to extract multi-element information (temperature, vegetation index, surface deformation, brightness) in the fire zones. Secondly, we constructed feature samples of concealed fire zones, open fire zones and non-fire zones, and built a AdaBoost_RF_StBP model based on the ensemble concepts of Boosting, Bagging, and Stacking. Finally, we inputted the fire zone samples into the AdaBoost_RF_StBP model to identify the fire zone range, and conducted a thorough analysis to verify the model's accuracy and generalizability. The specific technical route is shown in Fig. 1.

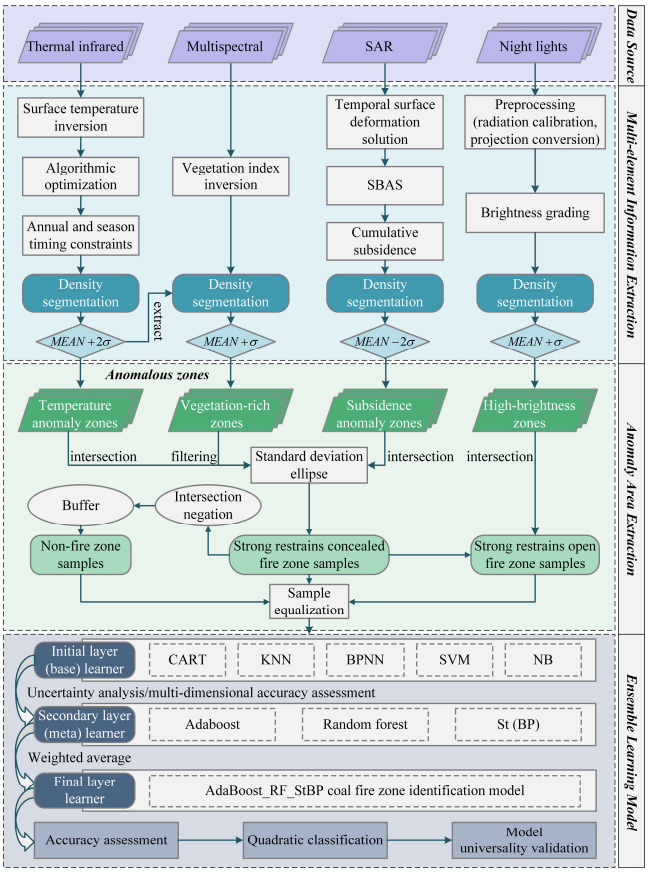

Fig. 1. Flow chart of the overall technical route of the method proposed in this study.

A. Extraction of Multi-element Information from Fire Zones Based on Multi-modal Remote Sensing Data

1) Inversion of Surface Temperature and Extraction of Temperature Anomalies: Firstly, we used the atmospheric correction (LaSRC) algorithm based on the thermal infrared radiation transfer equation and the single-window (SMW) algorithm proposed by Qin et al. [37] to invert the surface temperature. The calculation formula of the thermal infrared radiation transfer equation is expressed as.

$$L_{sen} = B(T_{sen}) = (\varepsilon B(T_s) + (1-\varepsilon)L_d) + L_u \quad (1)$$

where L_{sen} represents the on-board radiance ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$); B stands for the Planck function; T_{sen} is the on-board brightness temperature (K); ε is the surface emissivity; τ is the atmospheric transmittance; T_s is the surface temperature (K); L_d and L_u are the downward and upward atmospheric radiation ($\mathrm{W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}}$), respectively [38], [39].

The calculation formula of the single-window algorithm is.

$$T\_s = 1/C\ (a(1-C-D) + (b(1-C-D) + C + D)\ T\_sen - DT\_a) \quad (2)$$

where a and b are fixed coefficients of Planck equation; T_a is the average effective temperature of atmosphere (K); C and D are intermediate parameters related to surface emissivity; other parameters retain their previously mentioned meanings.

Existing methods for extracting temperature anomaly zones can be classified into two categories: the adaptive gradient threshold method and the manual threshold method. The former assumes that the surrounding overburden of the fire area has similar thermal conductivity and that the thermal energy decreases sharply from the burning center to the boundary of the fire area. Therefore, the average temperature corresponding to the temperature gradient extremum is used as the threshold to separate the anomaly area from the background area. The latter is based on the statistical principle that the surface temperature roughly follows a normal distribution. This method uses the mean ($\bar{x}$) of inverted surface temperature and different standard deviations (σ) to determine the segmentation threshold [28]. Given that coal fire zones typically exhibit higher surface temperatures compared to the average areas, we opted for the manual threshold method. This approach not only enhances the accuracy in identifying temperature anomaly zones but also reduces interference from areas with moderate to low temperatures. According to application requirements, we set the confidence level at 95%, i.e., took $\bar{x} + 2\sigma$ as the best segmentation threshold, and regarded areas above the threshold as initial temperature anomaly zones.

Factors such as feature type, climatic conditions, coal fire burning status, and temperature inversion algorithm at the time of remote sensing image imaging can affect the surface temperature inversion results in fire areas [40]. In addition, coal fire areas exist in three states: burned, burning and unburned. Thermal infrared data cannot effectively identify information about burned and cooled areas. To ensure high reliability of subsequent sample features, we intersected adjacent images to each winter season based on the optimal segmentation threshold. This mitigated the impact of surface temperature anomaly areas caused by direct solar radiation. We then superimposed them on an annual basis to attenuate the influence of climate and seasonal factors on surface temperature anomaly areas. This step also reduced the probability of coal fires being extinguished due to extreme weather, obtaining the final temperature anomaly zones.

2) Calculation of Vegetation Index and Extraction of Vegetation-rich Areas: The toxic substances released by coal fires greatly inhibit the normal growth of surrounding vegetation. Therefore, filtering out relatively vegetation-rich areas has an auxiliary effect on identifying fire zones. The Normalized Difference Vegetation Index (NDVI) effectively captures vegetation growth status and distribution information and can serve as a basis for extracting vegetation-rich areas. When using NDVI to filter out relatively vegetation-rich areas, it is necessary to determine a segmentation threshold. Unlike thermal anomaly zones, vegetation distribution varies around different fire zones. The average NDVI value of an entire image is often higher than that of fire zones, which is difficult to filter out if the segmentation is done directly by adding the mean and standard deviation. Therefore, we restricted the extraction of vegetation-rich areas within the final temperature anomaly zones. We experimentally adjusted the multiplicative relationship between the mean and standard deviation and combined it with the actual NDVI values to determine the segmentation threshold. Finally, areas exceeding this threshold were identified as vegetation-rich areas.

3) Time-series Surface Deformation Calculation and Extraction of Subsidence Anomalies: Continuous burning of coal fires can cause cracks in surrounding rock formations and cause surface subsidence. The development of InSAR/multi-temporal InSAR techniques allow for rapid and accurate probing of deformation information, offering significant potential in fire zone identification. We performed the SBAS-InSAR method proposed by Berardino et al. [41] to obtain the deformation time series over the coal fire areas. The basic principle is to assume that N view SAR images of the same area are obtained at times (t_0, t_1, ···, t_n), and small baseline sets below the temporal-spatial baseline threshold are selected to form M interferograms, then $N/2 \le M \le [N(N-1)]/2$. By using precise orbit (POD) with the reference digital elevation model (DEM) data, the interferograms are registered, multi-viewed and the differential interferograms are obtained after removing the flatland and terrain phases [42]. If the i-th differential interferogram is obtained at times t_A and t_B ($t_A < t_B$), the phase at a high coherence point x can be expressed as:

$$\begin{aligned}\Delta\phi_i(x,r) &= \phi(t_B,x,r) - \phi(t_A,x,r)\\ &\approx [\phi_d(t_B,x,r) - \phi_d(t_A,x,r)] + \frac{4\pi}{\lambda}\frac{B_\perp \Delta z}{r\sin\theta}\\ &\quad + [\phi_{atm}(t_B,x,r) - \phi_{atm}(t_A,x,r)] + \Delta n_i(x,r)\end{aligned} \quad (3)$$

where (x,r) is the coordinate of the high coherence point; $\phi(t_B,x,r)$, $\phi(t_A,x,r)$ are phase values of two SAR images; $\phi_d(t_B,x,r)$ and $\phi_d(t_A,x,r)$ are cumulative deformation phases in radar line-of-sight (LOS) direction relative to initial reference time t_0; $\phi_{atm}(t_B,x,r)$ and $\phi_{atm}(t_A,x,r)$ are atmospheric delay phases of two images; $\Delta n_i(x,r)$ is phase caused by noise; r and θ are slant range and incidence angle in line-of-sight direction; λ is radar wavelength.

Using a linear model to estimate the deformation phase of N view images is as follows:

$$\phi = (A^T A)^{-1} A^T \Delta\phi \quad (4)$$

where $A[M \times N]$ is a coefficient matrix corresponding to M differential interferograms and N SAR images at different times; ϕ is unknown deformation phase matrix of N SAR images at different times; $\Delta\phi$ is phase value matrix of M differential interferograms.

Since A^TA is often a singular matrix, the singular value decomposition methods are used in each subset to find least squares solutions of the unknown parameters in the sense of minimal norms, yielding linear deformation quantities and elevation errors. On the basis of the linear model, temporal and spatial filtering of the residual phase can further separate the atmospheric phase and the nonlinear deformation phase. Then, the final deformation can be obtained by combining the linear and nonlinear deformations [43].

In the actual process, we only performed a series of systematic treatments on the selected interferograms with good overall coherence. External atmospheric correction data was employed to assist in removing the atmospheric delay phase. We utilized the Delaunay triangulation, a network formed by irregular triangles, for 3D phase unwrapping, which notably increased the accuracy of unwrapping in regions with low coherence [44]. Thereby the final average deformation rate was obtained. Since the average deformation rate of all pixels roughly follows a normal distribution, we extracted the subsidence anomaly areas following the same approach used for extracting the initial temperature anomaly areas. Specifically, the optimal segmentation threshold was set by subtracting two times the standard deviation from its mean, and areas below this threshold were identified as subsidence anomaly zones.

4) Night Light Data Processing and Extraction of High-brightness Zones: Shallow floating coal has more contact with oxygen and is prone to open fires. This distinguishing feature sets fire areas apart from other ground objects. Therefore, night light remote sensing images can be used to detect open fire locations and assist in identifying fire areas.

Different ground object categories exhibit significant variations in pixel brightness values [45], [46], and the distribution range of brightness also varies significantly across the imagery [47]. We counted the brightness values of the entire night light image elements after the radiometric calibration and subtracted all areas with a value of 0, which were defined as bright zones. The bright zones contained fire areas, mining areas and a large number of scattered spots, which were further processed by the following steps in order to segment the feature types:

Step 1. Intersect the bright areas from two night light images obtained within the same month. Generally speaking, the burning conditions in fire areas are relatively stable, and open fire spots will not undergo significant changes on a large scale within a short period of time. Therefore, intersection processing can filter out scattered spots caused by sensors, partial high-radiation ground units and atmospheric scattering.

Step 2. Segment the intersected area into blocks and calculate the average brightness from both images. Segment the brightness values from the lowest to the highest into eight equal intervals. Then, combine temperature anomaly zones with subsidence anomaly zones to distinguish fire zones from mining areas.

Step 3. Calculate the brightness mean (x) and standard deviation (σ) for open fire spots. Due to the lower resolution of night light imagery, the light overflow effect might cause scattered open fire spots to be counted within a single pixel. Therefore, $\bar{x}+\sigma$ is used as threshold for density segmentation to further determine open fire position. Areas with values higher than this threshold are regarded as high-brightness zones.

B. Strongly Constrained Concealed Fire Zones, Open Fire Zones and Non-fire Zones Sample Delineation

To address the issue of limited overlap between the temperature anomaly zones and deep subsidence zones in the fire area, we created a 1 times standard ellipse based on the intersection of the temperature and subsidence anomaly areas. This serves as a strong constraint for the concealed fire zone sample range. The average center of the ellipse represents the spatial centroid of the fire zone's distribution. The long semi-axis represents the direction of fire area distribution. The short semi-axis represents the approximate extant of fire area distribution [48]. The strongly constrained open fire zones were overlaid on the strongly constrained concealed fire zones by adding the high-brightness zones. We assumed that all open fire zones were within the range of concealed fire zones. Non-fire zone samples were identified from two aspects: (i) By extracting independent temperature and subsidence anomaly zones and designating them as non-fire zones, the capability of the learner to distinguish anomaly zones characterized by these two factors as non-coal fire zones is enhanced, thereby further increasing the sensitivity of the learner; (ii) Extracting spots with no abnormal features. These spots were near mean in all four data sources, as supplement to non-fire zone samples. Therefore, the method for delineating non-fire area samples goes as follows: Using the standard deviation ellipse from the temperature and subsidence anomaly zones, we created a 100m buffer zone around it. The threshold of temperature anomaly and subsidence anomaly was lowered to 1 times standard deviation, and the independent temperature anomaly and subsidence anomaly data in the image were extracted (areas with no intersection between both). Data within the buffer was masked, and temperature and subsidence anomaly data were separated. Spots with no abnormal features were randomly extracted in the remaining area, and the number of spots was determined according to the image area. It was required to cover the entire test area.

In order to make each pixel value of multi-modal remote sensing data correspond to each other while increasing sample size, we firstly used the highest resolution data source as reference image and reprojected rest of multi-element information inversion results to same coordinate system. Secondly, we used nearest neighbor method to resample the resolution of all images. Finally, the concealed fire zones, open fire zones and non-fire zones samples were extracted using a mask for subsequent processing.

C. Construction of AdaBoost_RF_StBP Coal Fire Zone Identification Model

Different ensemble approaches produce various outcomes. Given their performance across diverse datasets, homogeneous ensemble methods like Boosting and Bagging necessitate substantial data to fully leverage their advantages [49]. Meanwhile, in the case of Stacking—a heterogeneous ensemble—the training data for its meta-learner is derived from the predictions of the base-learner, leading to potential error accumulation [50].

We combined both homogeneous and heterogeneous ensemble approaches, employing five models as base (initial layer) learners: Classification and Regression Tree (CART), K-Nearest Neighbor (KNN), Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), and Naive Bayes (NB). Adaptive Boosting (AdaBoost), Random Forest (RF), Stacked-BP Neural Network (StBP) were used

as meta (secondary layer) learners. Lastly, we derived the final layer of the AdaBoost_RF_StBP coal fire zone identification model by computing the weighted average of the identification probabilities from the meta-learners. Furthermore, we have used a genetic algorithm to automatically optimize the numerical and character-type hyperparameters involved in the base learners of ensemble learning. The conceptual framework for model construction is illustrated in Fig. 2. The parameter settings and construction steps of the AdaBoost_RF_StBP coal fire zone identification model are shown in Table I.

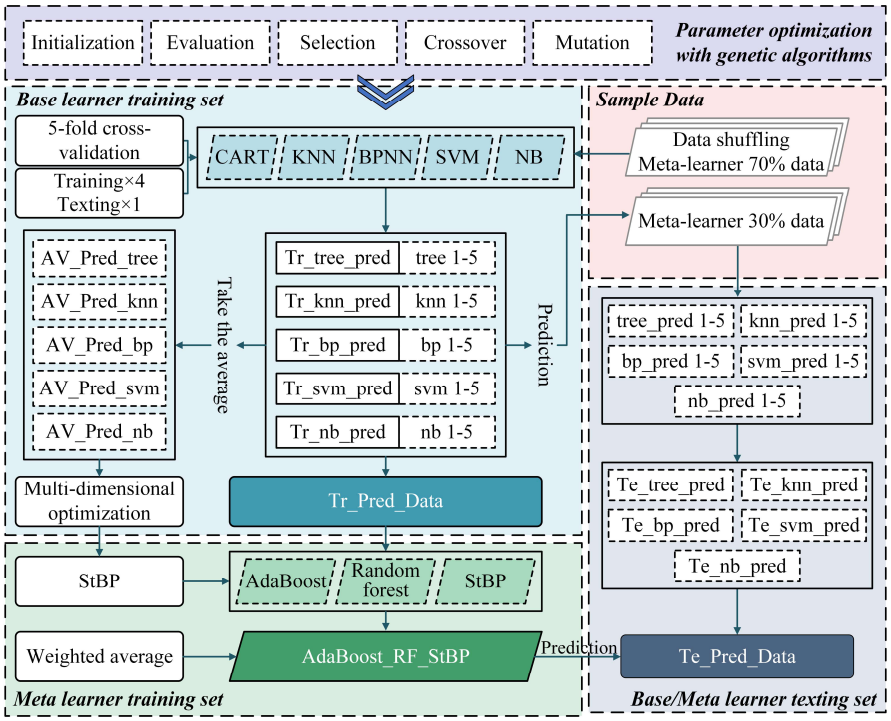

Fig. 2. The structure of AdaBoost_RF_StBP coal fire zone identification model.

TABLE I
BASE/META LEARNER MODEL PARAMETER SETTINGS

Base-learners (attribution)	Parameter	Value setting	Meta-learners (attribution)	Parameter	Value setting
CART (Tree)	Feature division point selection criteria	random	AdaBoost (Additive)	Number of base classifiers	100
	Maximum number of leaf nodes	50		Learning rate	1.0
	Maximum tree depth	10	RF (Tree)	Number of decision trees	100
	Internal node splitting minimum number of samples	2		Maximum number of leaf nodes	50
KNN (Neighbors)	Number of neighbors	8		Maximum tree depth	10
	Vector distance algorithm	Euclidean distance		Internal node splitting minimum number of samples	2
BPNN (Neural Network)	Activation function	relu	StBP (Neural Network)	Activation function	relu
	Learning rate / L2 regularization term	0.10/1.0		Learning rate / L2 regularization term	0.20/1.5
	Number of iterations	1000		Number of iterations	1000
	Hide layers / The number of neurons in each hidden layer	3/100		Hide layers / The number of neurons in each hidden layer	3/100
SVM (Kernel)	Penalty factor	1	Note: The StBP is a posterior model and is placed in this table in advance in a uniform parameter format.		
	Kernel functions	rbf			
	Error convergence conditions / Number of iterations	0.001/1000			
NB (Linear)	Prior distribution	Gaussian distribution			

1) Model Construction: The specific steps are as follows.

Step 1. Conduct ADASYN oversampling on the three sample data types outlined earlier to ensure balance. Subsequently, shuffle the dataset, allocating 70% for training and 30% for testing.

Step 2. Assume that the training data set is Tr and the test data set is Te. Use Tr to train the base-learners separately, and use the trained five models to predict Tr and Te, respectively. Use 5-fold cross-validation to reduce the risk of overfitting. After 10 times independent runs, Tr_tree_pred, Tr_knn_pred, Tr_bp_pred, Tr_svm_pred, Tr_nb_pred and Te_tree_pred, Te_knn_pred, Te_bp_pred, Te_svm_pred, Te_nb_pred are obtained. Take the average of the ten output results predicted by five models for Tr respectively to obtain

AV_Pred_tree, AV_Pred_knn, AV_Pred_bp, AV_Pred_svm, AV_Pred_nb. Statistical analysis of classification accuracy of base-learners in 10 trainings. Select the base-learner model with the highest accuracy to serve as the meta-learner model, denoted as St(model).

Step 3. Set Tr_Pred_Data = (Tr_tree_pred, Tr_knn_pred, Tr_bp_pred, Tr_svm_pred, Tr_nb_pred), Te_Pred_Data = (Te_tree_pred, Te_knn_pred, Te_bp_pred, Te_svm_pred, Te_nb_pred), use Tr_Pred_Data to train three meta-learners AdaBoost, RF and St(model) at secondary level. Use trained models to predict Te_Pred_Data to obtain prediction results of meta-learners.

Step 4. Calculate the weighted average from the three meta-learners to establish the final AdaBoost_RF_StBP coal fire zone identification model. Use Kappa coefficient to determine the weight. The larger the Kappa coefficient, the larger the corresponding weight, otherwise, the reverse. The weighting formula is.

$$W_{1,2,3} = \frac{Kappa_{1,2,3}}{Kappa_1 + Kappa_2 + Kappa_3} \tag{6}$$

where $Kappa_{1,2,3}$ represents the Kappa coefficient for a specific model. $Kappa_1$, $Kappa_2$, $Kappa_3$ represent the kappa coefficients of the three meta-learners, respectively.

To determine the final classification label predicted probability, multiply the weights of each model by the predicted probabilities of the three class labels naked_fire (nf), concealed_fire (cf), and unfire (uf), respectively. Formula for calculation is.

$$P_{nf} = \frac{P_n \times W_{1,2,3}}{W_1 + W_2 + W_3}, \quad P_{cf} = \frac{P_c \times W_{1,2,3}}{W_1 + W_2 + W_3}, \quad P_{uf} = \frac{P_u \times W_{1,2,3}}{W_1 + W_2 + W_3} \tag{7}$$

where P_{nf}, P_{cf}, P_{uf} represent the probability of open fire zones, concealed fire zones and non-fire zones after weighted average of the three meta-learners, respectively; P_n, P_c, P_u represent the probability of open fire zones, concealed fire zones and non-fire zones predicted by each model, respectively; $W_{1,2,3}$ represents the weight of a specific model; W_1, W_2, W_3 represent the weights of the three meta-learners, respectively.

2) Genetic Algorithm Construction: The specific steps are as follows.

Step 1. Define the fitness function: This function accepts the creation functions of five base learners, a dictionary of hyperparameters, and training data. We use accuracy as the metric, evaluate the performance of each classifier using cross-validation, and return the mean accuracy of each learner over five runs as the individual's fitness.

Step 2. Initialize the population: If the hyperparameter is character-type, a value is randomly selected from the numerical mapping list. If the hyperparameter is an integer with an upper bound, an integer is randomly generated between its lower and upper bounds. If there is no upper bound, an integer is randomly generated between its lower bound and ten times its lower bound. If the hyperparameter is a floating-point type, its random number generation strategy is the same as that of the integer type, but the random number is a floating-point number. Subsequently, each generation of the population contains 50 individuals.

Step 3. Selection: The tournament algorithm is used, with the number of competitors in each tournament being four. Individuals with high fitness are selected from the population.

Step 4. Crossover: This function accepts two parent individuals and a dictionary of hyperparameters, and generates a child individual.

Step 5. Mutation: Set the mutation probability for each offspring individual to 10% to enhance randomness.

Step 6. Iteration: Execute the genetic algorithm with the number of iterations set to 20.

III. Test Areas

A. Overview of Test Area

The eastern part of the Junggar Basin in Xinjiang is rich in coal resource. This region exhibits a continental desert climate, which is characterized by arid to semi-arid conditions, minimal rainfall, significant annual and daily temperature fluctuations, and intense solar radiation. Severe weathering conditions of soil and rocks result in widespread distribution of coal fire areas. In this study, we focused on three pivotal coalfield fire regions: Jiangjun Gobi, Jiangjun Temple in Qitai, and Sandaoba in Miquan—all situated in the eastern segment of the Junggar Basin. For a more thorough analysis, the Beishan coal mines region is also included. Fig. 3 shows the precise geographical locations.

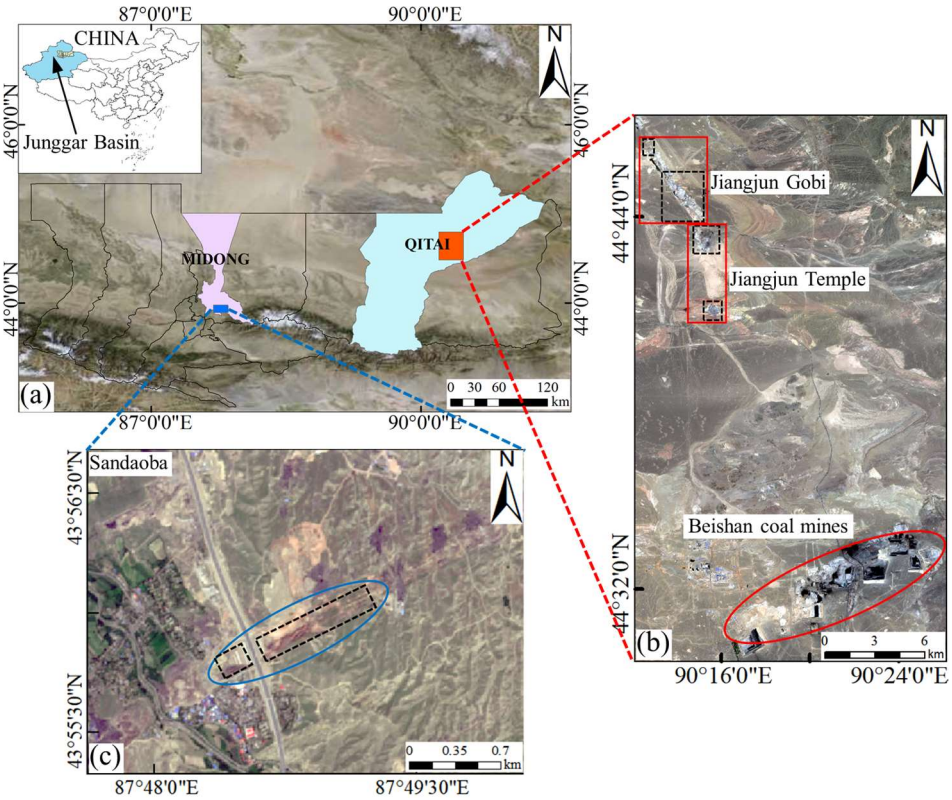

Fig. 3. Geographical location of the test areas. (a) The location of Qitai County and Midong District in the eastern Junggar Basin of China. (b) The red boxes from north to south represent the Jiangjun Gobi fire zone, the Jiangjun Temple fire zone, and the Beishan coal mines area, respectively. The black box represents the northern and southern parts of the fire zone. (c) The blue box represents the Sandaoba fire zone, and the black box represents the eastern and western parts of the fire zone.

The Jiangjun Gobi fire area is located in the northeast of Qitai County, Changji Hui Autonomous Prefecture, Xinjiang. Its central coordinates are 90°14'40" E and 44°45'9" N. This fire area consists of two sections: The northern section measures an average 341m in length and 148.7m in width, encompassing approximately 50,679m^2. The southern section extends 2,974m in length with an average width of 238m, covering around 708,909m^2. The Jiangjun Temple fire area is located south of the Jiangjun Gobi. The coordinates of its center are 90°42'15" E and 44°27'52" N. Its two primary segments are: The northen fire zone is arranged along the coal seam outcrop from northwest to southeast, with a total area of 79,880m^2. The southern fire zone is dispersed along the coal seam outcrop in an east-west orientation, with a total area of 40,633m^2. Lastly, the Sandaoba fire area is located in the east of Jianquan Gully in Midong District, Urumqi City. It is centered at 87°25'27" E and 43°47'41" N. This region also comprises two sections: The eastern section, shaped like a long strip running from the southwest to northeast, measures around 2,695m in length and varies between 116m and 257m in width. It spans roughly 564,644m^2. The western section is irregular in form, with a length of about 145m from east to west and a width of about 100m from north to south, covering an area of 9,419m^2.

B. Data Sources

Four types of remote sensing data are adopted in this study: Landsat-8 multispectral (OLI) B4, B5 band data; Thermal infrared (TIRS) B10 band data; Sentinel-2A multispectral (MSI) B4 and B8 band data; C-band Sentinel-1A ascending data in IW working mode with VV polarization and LuoJia-1 GEC system geometric correction product. External data comprise: DEM data with a 30m resolution from the Shuttle Radar Topography Mission (SRTM); the Generic Atmospheric Correction Online Service for InSAR (GACOS) offering a spatial resolution of 90m and a temporal resolution of one minute; and detailed survey reports of the three fire zones from the Xinjiang Coalfield Fire Extinguishing Engineering Bureau. The acquisition timings, quantities and resolutions of the remote sensing data for the test area are presented in Table II.

TABLE II
VARIOUS IMAGE DATA INFORMATION IN THE TEST AREA

Parameters / Sensors	Jiangjun Gobi, Jiangjun Temple				Sandaoba			
	Time		Number of images	Resolution(m)	Time		Number of images	Resolution(m)
Sentinel-1A	20170321	20200317	36	5×20	20150124	20170125	29	5×20
Sentinel-2A	20170610 20180615	20190521	3	10	20160804	20170610	2	10
Landsat-8	20171028 20171129 20181031	20181116 20191119 20200122	6	OLI:30 TIRS:100	20151030	20161101	2	OLI:30 TIRS:100
LuoJia-1	20181009	20181029	2	130	—		0	130

IV. Results

A. Multi-modal Remote Sensing Data Inversion and Anomaly Area Extraction

1) Surface Temperature Inversion and Temperature Anomaly Identification: Using the LaSRC and SMW algorithms in tandem with six Landsat-8 image scenes, we conducted surface temperature inversions for the test areas of Jiangjun Gobi and Jiangjun Temple. The statistical histograms of these inversion results, along with their mean and standard deviation are displayed in Fig. 4.

Analyzing Fig 4 reveals that the temperature curves for the first five periods roughly follow a normal distribution. Simultaneously, due to a significant snow cover in the sixth period, the temperature trend significantly shifted towards cooler values. The inversion outcomes from the two algorithms are comparable, but the LaSRC algorithm demonstrates a smaller standard deviation compared to the SMW, indicating a lower degree of dispersion and stronger stability. Therefore, we chose the surface temperature data derived from the LaSRC algorithm to delineate the initial temperature anomaly areas. These data were then used in subsequent time-series overlay procedures.

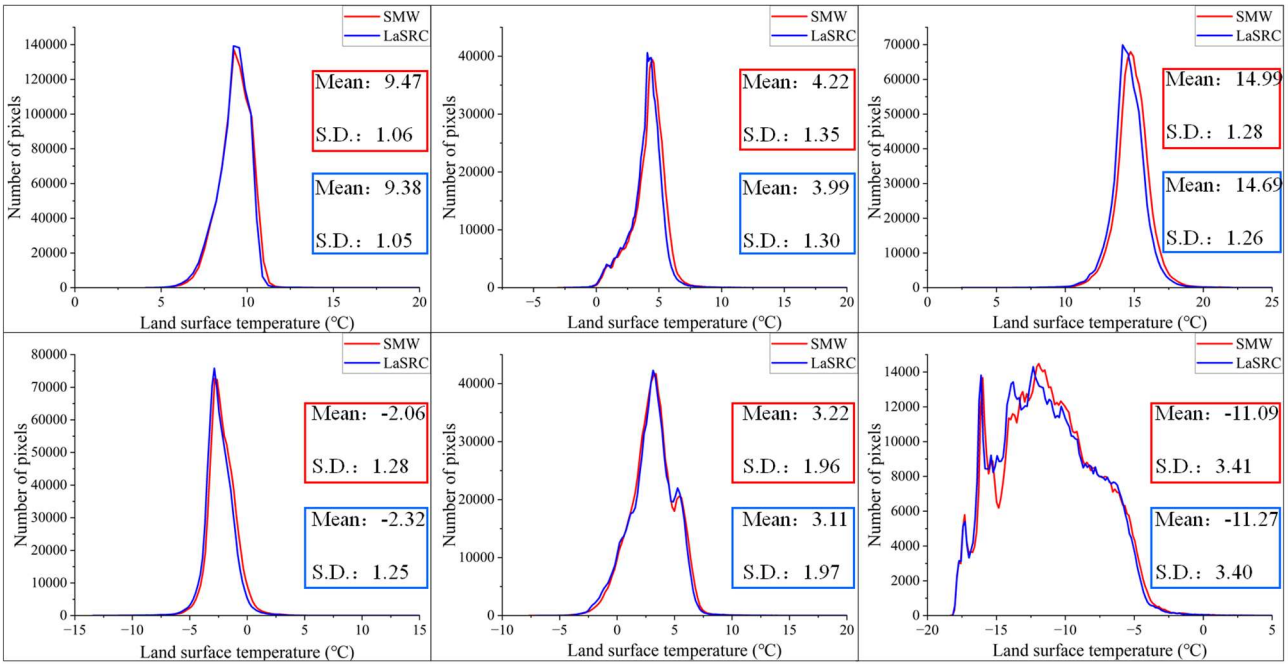

Fig. 4. Land surface temperature inversion statistical chart of LaSRC/SMW algorithm (The temperature curves, as well as the mean and standard deviation of the inversion results for the LaSRC algorithm and the SMW algorithm, are denoted in red and blue, respectively.).

Fig. 5(a)-(f) depicts the initial temperature anomaly areas across the six periods post-threshold segmentation. It can be observed that there are significant variations in the range of temperature anomaly areas across different seasons within the same year and between different years. When using a single temperature anomaly result, there may be problems with the range being "overestimated" or "underestimated". Fig. 5(g) shows the final temperature anomaly areas under annual and seasonal sequential overlay limitations. Notably, these temperature anomaly zones are predominantly situated in the southern fire area of Jiangjun Gobi, the northen fire area of Jiangjun Temple, and the Beishan coal mines area. Independent temperature anomalies have been efficiently eliminated. Additionally, it's evident that solely using temperature anomalies is insufficient to distinguish between fire areas and mining areas. Therefore, it is essential to incorporate additional information for further constraints.

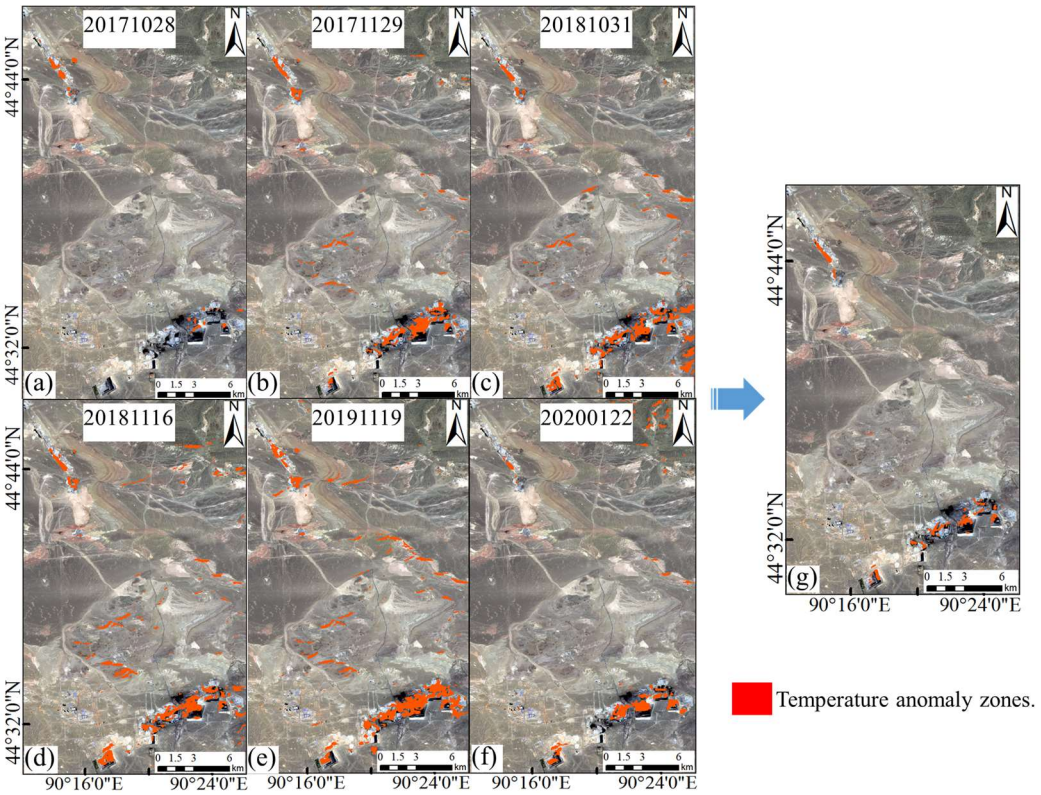

Fig. 5. The extent of the surface temperature anomaly zones. (a)-(f) are the initial temperature anomaly areas of 6 scenes, where the red part represents the extent of temperature anomalies. (g) The final temperature anomaly area after being constrained by annual and seasonal sequences.

2) Vegetation Index Inversion and Vegetation-rich Area Identification: We conducted radiometric calibration, atmospheric correction, and topographic correction uniformly on three scenes of Sentinel-2A imagery, then calculated the NDVI. As illustrated in Fig. 6(a)-(c), the NDVI in the test area is generally low. However, higher values appear in the northern Jiangjun Gobi, Jiangjun Temple, and southern Beishan coal mines. This observation is counterintuitive, as fire and mining activities typically inhibit vegetation growth. Therefore, by averaging the results of NDVI calculations over three periods, we extracted and quantified the mean and standard deviation of NDVI within the final temperature anomaly areas. The experiment found that using the mean plus one standard deviation as the segmentation threshold not only retained most of the final temperature anomaly areas, but also effectively filtered out areas with high NDVI [Fig. 6(d) and (e)]. By filtering these areas from the temperature anomaly zones in subsequent processing, the error rate of delineating temperature anomalies can be reduced, further enhancing the reliability of the fire zone samples.

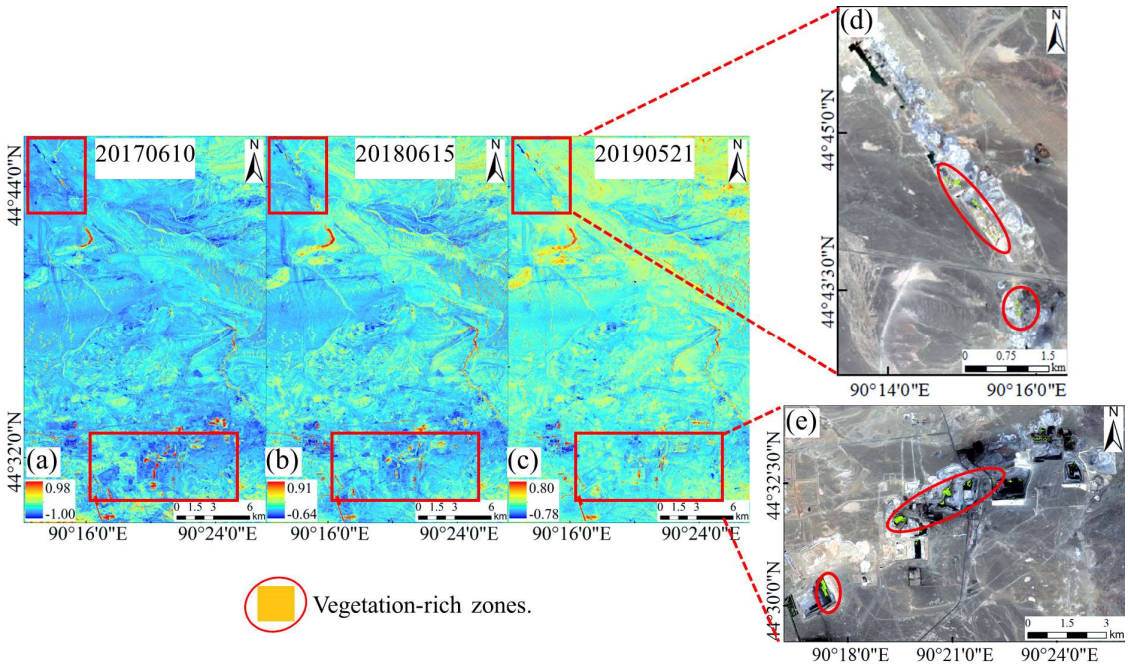

Fig. 6. NDVI and vegetation-rich zones. (a)-(c) are the NDVI inversion results of 3 scenes. The red boxes at the top represent the northern fire zones of Jiangjun Gobi and Jiangjun Temple, and the red boxes at the bottom represent the Beishan coal mines area. (d) and (e) The green part inside the red ellipse represents the area with vegetation-rich zones.

3) Time-series Surface Deformation Detection and Subsidence Anomaly Identification: For SAR image processing, we utilized the Sarscape software on 36-view SAR images, and then trimmed the images to fit the test area dimensions. Given that the test area mostly comprises the Gobi with sparse vegetation, we adjusted the thresholds to balance temporal coherence and processing efficiency: setting the maximum spatial baseline threshold to 2% of the critical baseline, the maximum temporal baseline threshold to 120d, and the minimum time-space baseline thresholds to 0. Finally, we obtained a total of 114 pairs of small baseline connection diagrams. Spatial and temporal baseline distribution is shown in Fig. 7.

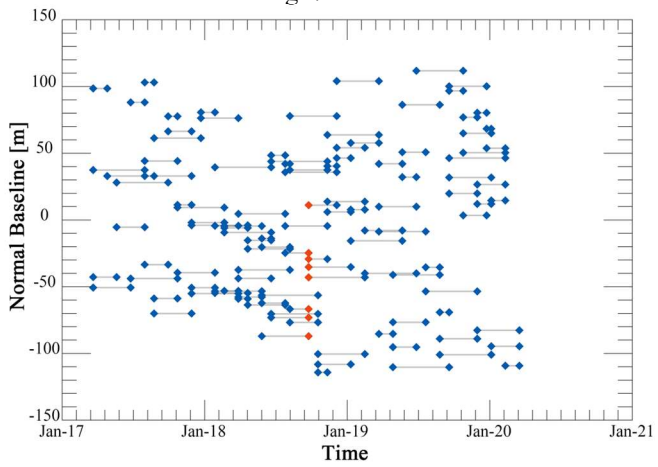

Fig. 7. Spatial-temporal baseline distribution.

Following the above data processing approach, we first aligned the images and applied 4:1 multi-viewing in both distance and azimuth directions. We then removed topographic and orbital phases from the interferometric phases by using the 30m resolution SRTM DEM data and precision orbit data (POD), respectively.

We applied the Delaunay Minimum Cost Flow (Delaunay MCF) method algorithm [51] for phase unwrapping and used GACOS data to correct the atmospheric delay phase. This allowed us to estimate the average deformation rate and the residual phase component. As a result, we obtained the average deformation rate for the Jiangjun Gobi and Jiangjun Temple fire areas from March 21, 2017 to March 17, 2020 [Fig. 8(a)-(c)].

Observations reveal that the overall deformation in the test areas are stable. Notable subsidence is primarily evident in the Jiangjun Gobi, Jiangjun Temple fire area, and Beishan coal mines area. The average subsidence rate in the fire areas is over 30 mm/a, with the southern portion of the Jiangjun Gobi fire area reaching up to 60 mm/a. Meanwhile, the majority of the Beishan coal mines demonstrate an average subsidence rate of 60 mm/a, indicating that it is difficult to distinguish between fire areas and mining areas using deformation information alone. Fig. 8(d) shows the subsidence anomaly areas obtained after threshold segmentation, mainly covering the eastern excavation pit of the southern Jiangjun Gobi fire area, the northern fire area of Jiangjun Temple, and most parts of the Beishan coal mines.

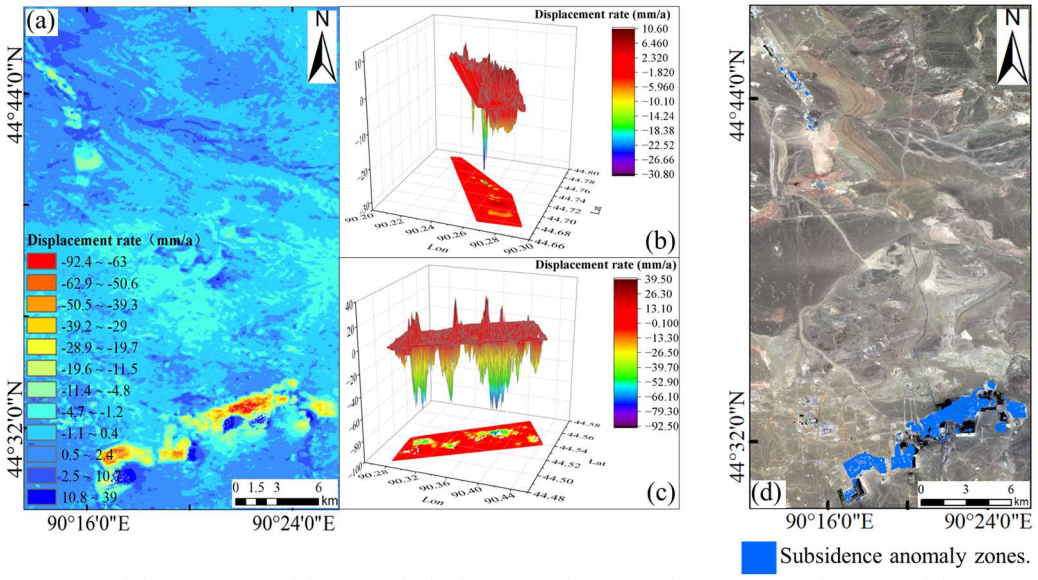

Fig. 8. The three-year average deformation rate and the range of subsidence anomaly zones in the test area. (a) The average deformation rate of the Jiangjun Gobi, Jiangjun Temple fire areas, Beishan coal mines and surrounding areas over three years. (b) and (c) are the three-dimensional mappings of the average deformation rates of the Jiangjun Gobi, Jiangjun Temple fire areas and Beishan coal mines area, respectively. (d) The range of the subsidence anomaly areas.

4) Night Light Data Brightness Analysis and High-brightness Zones Identificatgion: Fig. 9(a), (b), and (c) show brightness grading results for the fire and mining areas after intersecting two night light images. It can be observed that the fire zones have higher brightness values, with the maximum reaching 0.03920. In contrast, the mining area demonstrates a significantly subdued brightness, with a maximum value of 0.005320, which is only 1/7 of the brightness in the fire area. Fig. 9(b) and (c) show differences in brightness distribution. The fire area has a higher degree of brightness aggregation, with only one brightness peak in the figure. This may be due to the dispersion effect of the fire, causing open fire spots to gather in a pixel cell and leading to an accumulation of brightness values within entire cell. On the other hand, the mining area has a more dispersed brightness range, with two brightness peaks in the figure, and scattered smaller peaks around them. This may be due to frequent and widely distributed human activities in the mining areas, resulting in stable light brightness radiation.

Moreover, at the two brightness peaks in the mining areas, there are no anomalies in temperature or subsidence, suggesting that these locations might be fixed luminous structures. Therefore, through a comprehensive analysis encompassing temperature, subsidence, and brightness data,

we segmented the fire and mining areas, exclusively extracting the fire zone range [Fig. 9(d)]. The threshold was determined by adding the mean brightness value to twice its standard deviation, identifying the high-brightness zones [Fig. 9(e)]. This area serves as the open fire area.

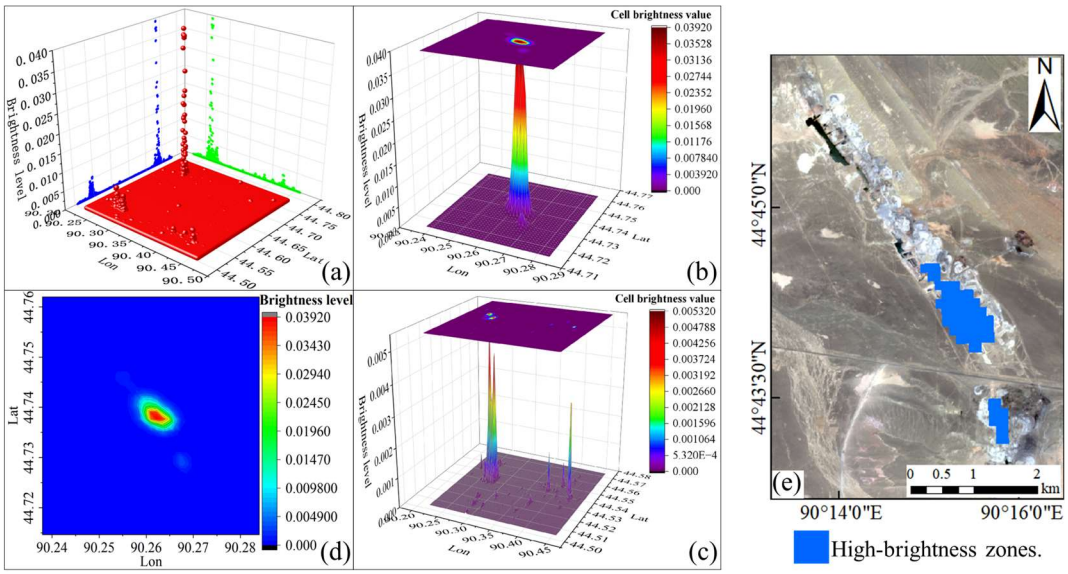

Fig. 9. Grading results of night light data and high-brightness zone range. (a) The brightness level of Jiangjun Gobi, Jiangjun Temple, Beishan coal mines and surrounding areas. (b) and (c) are the three-dimensional mappings of the brightness levels of Jiangjun Gobi, Jiangjun Temple fire zones and Beishan coal mines, respectively. (d) The brightness grading results of Jiangjun Gobi and the northern fire zone of Jiangjun Temple. (e) The area of high-brightness.

B. Strongly Constrained Delineation Sample Results of Concealed Fire Zones, Open Fire Zones and Non-fire Zones

After segmenting the fire areas of Jiangjun Gobi and Jiangjun Temple, as well as the Beishan coal mines, we employed a standard deviation ellipse for the overlapping regions of temperature and subsidence anomalies primarily located in the southern Jiangjun Gobi fire area and the northern Jiangjun Temple fire area. Fig. 10(a), (b), and (c) illustrate the samples derived from the strongly constrained concealed fire zones, the open fire zones, and the non-fire zones, as established by the methods described.

In Fig. 10(a), the standard deviation ellipse, formed based on the intersection of temperature and subsidence anomaly regions, exhibits pronounced elongation and directional characteristics, mirroring the overarching pattern of the fire area. A single standard deviation ellipse effectively captures approximately 68% of the fire areas, striking an optimal balance between sample comprehensiveness and precision. Additionally, Fig. 10(a) further delineates the boundaries of the open fire zones, highlighting a concentration of open fire activity within the southern Jiangjun Gobi fire area. Meanwhile, Fig. 10(b) and (c) presents the non-fire zones, featuring three distinct temperature anomaly samples, a singular subsidence anomaly sample, and 2000 randomly generated anomaly-free data spots. The temperature anomaly areas are partly concentrated in the southern part of the mountain and partly outside the buffer area of the Beishan coal mines, which may be related to solar radiation and human activities on the ground, respectively. The Beishan coal mines subsidence anomaly is the most typical in the whole image, so it is used as a separated subsidence anomaly area.

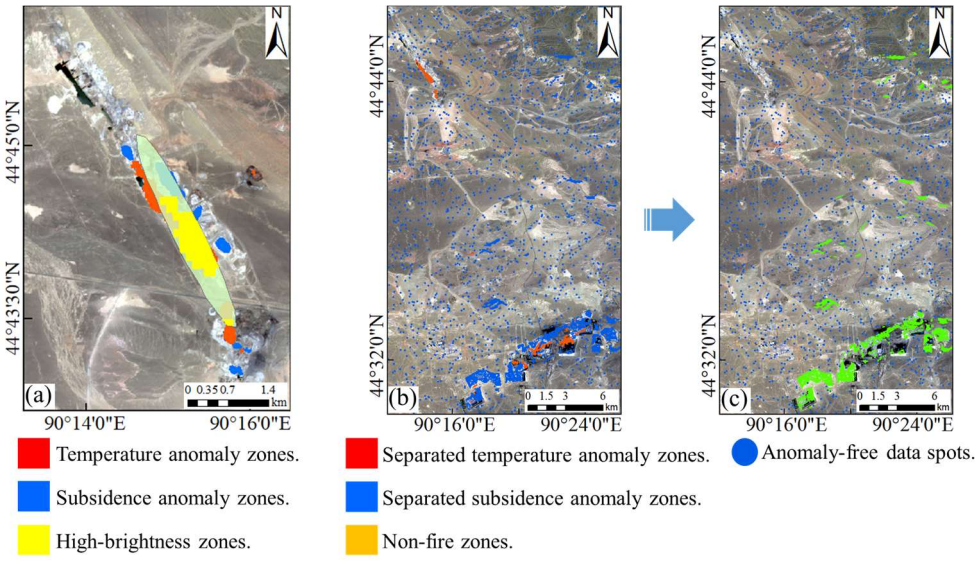

Fig. 10. Strongly constrained samples from concealed fire zones, open fire zones, and non-fire zones. (a) The strongly constrained concealed fire zones and the strongly constrained open fire zones, where the ellipse is the 1 times standard deviation ellipse. (b) The separated temperature and subsidence anomaly areas (The vegetation-rich areas have been filtered out). (c) The non-fire areas integrated from (b).

C. Results of the AdaBoost_RF_StBP Coal Fire Zone Identification Model

Considering the imaging times of the night light images and the quality of inversions for various types of images, we undertook a band fusion, integrating the average deformation rate, the vegetation inversion data (from June 15, 2018), the temperature inversion data (from October 31, 2018), and the night light inversion data (from October 9, 2018). More precisely, we took Sentinel-2A as the reference image while reprojecting the rest of the data, and up-sampling using the nearest neighbor method. Among them, the final temperature anomaly areas are obtained by annual and seasonal timing constraints. The vegetation-rich areas are obtained by taking the mean of two images in the summer of each year and performing threshold segmentation. The deformation anomaly areas are obtained by taking the temporal average deformation rate and performing threshold segmentation. The high-brightness areas are obtained through brightness grading and density segmentation. Subsequently, we extracted three distinct sample types, as detailed in above, and utilized them for training within the base-learner framework.

To evaluate the classification results' quality across multiple dimensions, we employed metrics including accuracy, recall, precision, F1_Score, Kappa coefficient, and confusion matrix heatmap. Generally speaking, in the context of imbalanced samples, accuracy might not be a reliable indicator to measure outcomes, even when ADASYN oversampling has been applied. Metrics such as recall, precision, and F1_Score emphasize the relationship between predicted positive samples and actual positive samples (i.e., the set of object labels of interest). Thus, the values of these metrics provide more insightful assessments of final classification precision.

1) Base-Learner Classification Accuracy Analysis: Table III presents the accuracy outcomes of the training phase for each model within the base learner set, while the heatmap of the confusion matrix for each model validation set test is given in Fig. 11.

TABLE III
BASE-LEARNER ACCURACY FOR EACH MODEL

Model (Data set)	Index	Accuracy	Recall	Precision	F1_score	Kappa
CART	Training set	0.808	0.808	0.826	0.790	0.710
	Cross-validation set	0.854	0.854	0.879	0.846	
KNN	Training set	0.873	0.873	0.896	0.868	0.781
	Cross-validation set	0.848	0.848	0.864	0.843	
BPNN	Training set	0.869	0.869	0.896	0.863	0.809
	Cross-validation set	0.872	0.872	0.899	0.866	
SVM	Training set	0.649	0.649	0.6	0.595	0.569
	Cross-validation set	0.629	0.629	0.591	0.588	
NB	Training set	0.780	0.780	0.798	0.765	0.671
	Cross-validation set	0.781	0.781	0.798	0.766	

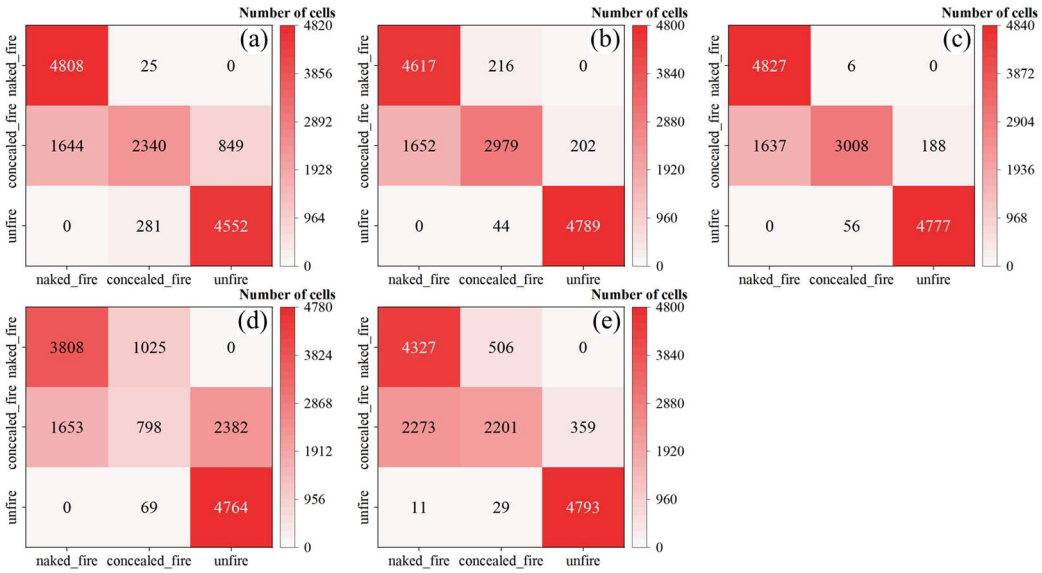

Fig. 11. Base-learner test data confusion matrix heatmap. (a)-(e) are CART, KNN, BPNN, SVM, and NB in order. (The horizontal axis represents the actual sample, and the vertical axis represents the predicted sample).

From the analysis of Table III, it can be observed that the KNN model exhibits the highest values for accuracy, recall, and F1_Score in the training set, recorded at 0.873, 0.873, and 0.868, respectively. Its precision is 0.896, on par with the BPNN model. Within the cross-validation set, the BPNN model outperforms the rest, boasting values of 0.872, 0.872, 0.899, and 0.866 for accuracy, recall, precision, and F1_Score, respectively. When these findings are combined with the Kappa correlation coefficient and insights from Fig. 11, it becomes evident that the BPNN model achieves the highest classification accuracy and correct prediction count. Consequently, we selected the BPNN model to serve as the meta-learner (StBP) for the computation involving secondary learners.

2) Meta-learner Classification Accuracy Analysis: The prediction results of the five base-learners were used as training data inputted for the meta-learner. Each learner still adopted a ten-fold cross-validation method to reduce the risk of overfitting, and was run independently ten times to output the training results for the meta-learner. Table IV presents

the accuracy of the training results for each model in the meta-learner. Fig. 12 provides the confusion matrix heatmaps for the validation set of each model, while Fig. 13 displays the box plot representing the stability of each model after ten independent runs.

TABLE IV
META-LEARNER ACCURACY FOR EACH MODEL

Model (Data set) \ Index		Accuracy	Recall	Precision	F1_score	Kappa
AdaBoost	Training set	0.836	0.836	0.851	0.841	0.737
	Cross-validation set	0.805	0.805	0.842	0.807	
RF	Training set	0.959	0.959	0.958	0.958	0.934
	Cross-validation set	0.942	0.942	0.943	0.940	
StBP	Training set	0.955	0.955	0.954	0.954	0.932
	Cross-validation set	0.944	0.944	0.945	0.943	

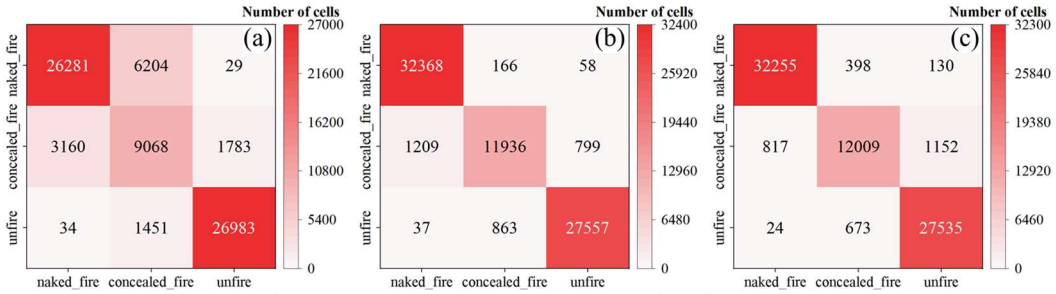

Fig. 12. Meta-learner test data confusion matrix heatmap. (a)-(e) are AdaBoost, RF and StBP in order.

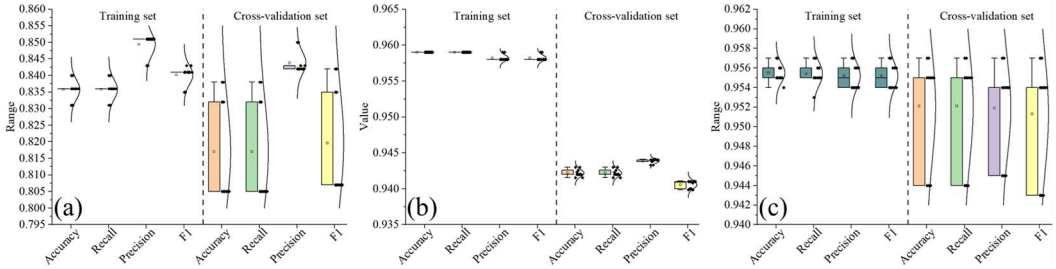

Fig. 13. Box plot of stability evaluation index of meta-learner models (The sequence is the same as in Fig. 13).

Based on the analysis of Table IV and Fig. 12, it can be observed that the overall estimation accuracy of the meta-learners has significantly improved compared to the base-learners. For the RF model in the training set, the accuracy, recall, precision, and F1_Score have increased by 0.09, 0.09, 0.062, and 0.095 respectively compared to the BPNN model in the base-learners. In cross-validation set, the improvements are 0.07, 0.07, 0.044, and 0.077, respectively. The performance of the StBP model is comparable to the RF model, with slightly higher metrics in the cross-validation set than the RF model. Performance of AdaBoost model is slightly worse, but still comparable to that of BPNN and performs better than both the SVM and NB learners.

Among the meta-learners, both RF and StBP achieve the Kappa coefficient of over 0.93, indicating high degree of consistency between actual sample labels and predicted sample labels. Further combined with Fig. 13, the boxplots of the various metrics for the RF model in both the training and cross-validation sets are highly flat, meaning the estimation results are nearly the same for each learning iteration, making it the most stable model. The AdaBoost model also demonstrates high stability in the training set, with some fluctuations observed in the cross-validation set. The StBP model demonstrates minor fluctuations as well.

From this, it is evident that integrating learners based on Bagging, Boosting, and Stacking algorithms can further effectively enhance sample classification accuracy. The Bagging and Stacking ensemble algorithms yield the best improvement results.

3) Analysis of Identification Results of AdaBoost_RF_StBP Coal Fire Zone Identification Model: For practical application, we applied a secondary mask to the Jiangjun Gobi and Jiangjun Temple fire areas within the synthesized band image. Subsequently, this data was fed into the AdaBoost_RF_StBP model for coal fire zone identification, aiming to pinpoint the ultimate fire area demarcation. For comparison, we input the masked data into AdaBoost, RF and StBP separately. This allows us to examine the distinctions between the individual use of Boosting, Bagging, and Stacking ensemble strategies versus their combined application. To validate the universality of the AdaBoost_RF_StBP coal fire zone identification model, we also took the Sandaoba fire area as the test area. The temperature, deformation, and vegetation index imagery were re-projected, re-sampled, and band-merged in the same manner. After masking to extract data and inputting it into the model, the identified fire zone regions was compared with the actual fire zone regions.

For a more intuitive comparative analysis, we enlarged the actual fire zone regions, as shown in Fig. 14 (a), (b), and (g), while (c)-(f) and (h) represent the corresponding model identification results. Table V lists the actual detected area and identified area for each fire zone.

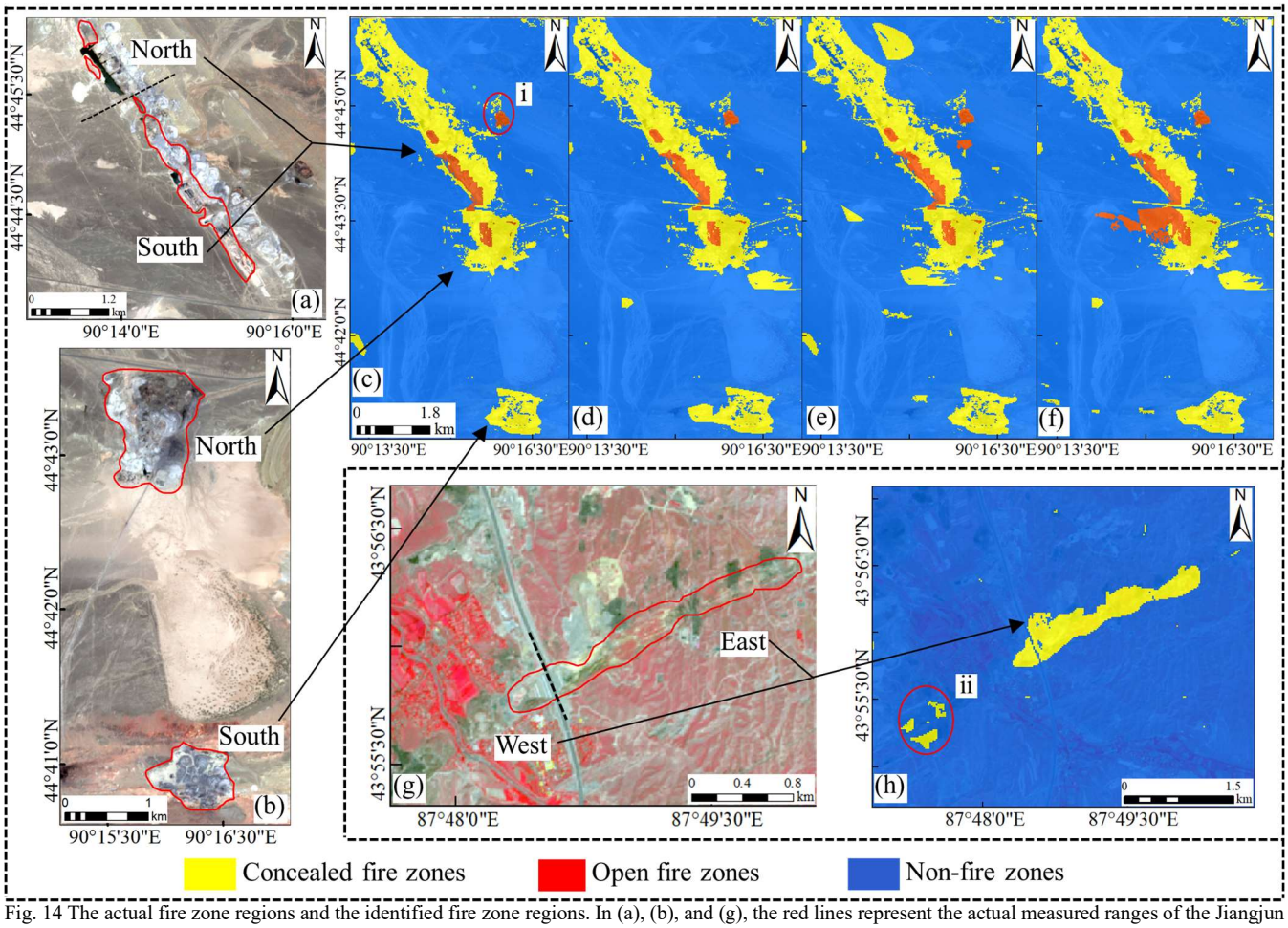

Fig. 14 The actual fire zone regions and the identified fire zone regions. In (a), (b), and (g), the red lines represent the actual measured ranges of the Jiangjun Gobi fire area, the Jiangjun Temple fire area, and the Sandaoba fire area, respectively. (c) to (f) represent the identification results in the Jiangjun Gobi and Jiangjun Temple fire areas after secondary masking by the AdaBoost_RF_StBP, AdaBoost, RF and StBP model, respectively. (e) presents the identification result for the Sandaoba fire area using the AdaBoost_RF_StBP model.

TABLE V
DETECTION AREAS OF FIELD FIRE AREAS AND IDENTIFICATION OF FIRE AREAS

Model \ Name of fire areas		Jiangjun Gobi (north and south)	Jiangjun Temple (north and south)	Sandaoba (west and east)
Actual fire zone areas		50679m^2+708909m^2	79880.05m^2+40633.18m^2	9419m^2+564644m^2
AdaBoost_RF_StBP	Identification area	932774m^2	132323m^2	665339m^2
	Misclassification area/percentage	173186m^2/22.8%	11810m^2/9.8%	91276m^2/15.9%
AdaBoost	Identification area	971513m^2	137626m^2	—
	Misclassification area/percentage	211925m^2/27.9%	17112m^2/14.2%	—
RF	Identification area	957840m^2	136059m^2	—
	Misclassification area/percentage	198252m^2/26.1%	15546m^2/12.9%	—
StBP	Identification area	977589m^2	138228m^2	—
	Misclassification area/percentage	218001m^2/28.7%	17715m^2/14.7%	—

As depicted in Fig. 14, it is evident that the fire zones identified by the model show a high degree of overlap in orientation with the actual surveyed fire zones. The concealed fire zones, open fire zones, and non-fire zones are well-separated. In conjunction with Table V, it is clear that the overall areas of the identified fire zone tend to be larger. Specifically, the recognition areas of AdaBoost_RF_StBP in Jiangjun Gobi, Jiangjun Temple and Sandaoba fire areas account for 122.7%, 109.8% and 115.8% of the actual fire areas, respectively. The misclassification rates are 22.8%, 9.8% and 15.9%, respectively. Extensive excavation pits surround the northern regions of both Jiangjun Gobi and Jiangjun Temple and predominantly contribute to the subsidence anomaly. The identified results in these two areas exhibit significant deviations from the actual measurements. Conversely, the southern part of Jiangjun Temple is relatively isolated, with comprehensive fire zone identification and minimal surrounding impurities. Moreover, region i exhibits an isolated misjudgment, being identified as having an open fire. This might be attributed to significant contributions from both temperature and subsidence in the area, or it might suggest an expansion of the fire zone following the actual survey. In the Sandaoba fire area, there is a significant misidentification at region ii. The identification results of the AdaBoost, RF, and StBP models show more fragmented fire areas, which is inconsistent with the trend that fire areas should be clustered. The StBP model identifies more open fire areas, but in combination with field survey results, this part should be concealed fire areas. The identification results of the three single models in Jiangjun Gobi and Jiangjun Temple are generally consistent with the AdaBoost_RF_StBP model. In addition, in conjunction with Table V, the misclassification rates of the three single models are high, exceeding 25% in

Jiangjun Gobi and 10% in Jiangjun Temple, which indirectly reflects the robustness of the AdaBoost_RF_StBP model.

The phenomenon of missed detection in the three fire areas is virtually non-existent, effectively compensating for omissions by filtering methods due to the small overlap range of temperature and subsidence anomalies in the vertical direction. Most of the misidentified areas are influenced by the surrounding excavation pits and subsidence anomalies. Yet, these excavation pits contain a significant amount of floating coal. If left unaddressed, it could elevate the potential for re-ignition and subsequently influence the trajectory of subsequent combustion spread. Additionally, model identification results include several open fire areas. These areas exhibit the most pronounced combined anomalies and warrant vigilant monitoring. In summation, utilizing multi-source remote sensing in tandem with ensemble learning method for fire area identification is not only viable but also yields relatively precise outcomes across expansive regions.

V. Discussion

A. Impact of Background Areas on Accuracy of Coal Fire Identification

After coal undergoes spontaneous combustion, it develops into a fully burning fire source. The high-temperature area propagates to the surrounding low-temperature area through heat conduction, forming destructive cracks under the action of thermal stress. After the coal body is burned, the overlying rock layer undergoes deformation and destruction to form a large range of collapse cracks when the external load energy and its internal energy lose balance. Therefore, the spatio-temporal differences between the coal fire area and the background area are mainly reflected in temperature and deformation, and indirectly affect elements such as vegetation and soil. By constraining in space and time, and utilizing multi-source remote sensing technology, the characteristics of the fire area can be highlighted in the entire background area, achieving the purpose of identification. During the sample extraction phase, in longstanding fire areas like Sandaoba [Fig. 15(a)], directly extracting the intersecting areas of temperature and subsidence anomalies as fire zone samples is more representative than constructing a standard deviation ellipse. Table V also reflects that the identification accuracy for the Sandaoba fire area is consistently higher compared to the Jiangjun Gobi and Jiangjun Temple fire areas. One reason for this is that actual fire areas are often surrounded by mining areas. The area delineated by the standard deviation ellipse for the fire areas sometimes resembles these mining areas. For fire zones with shorter burning times, such as Jiangjun Gobi and Jiangjun Temple illustrated in Fig. 15(b), the overlap of various anomalies is not evident, leading to an insufficient collection of fire zone samples. While the standard deviation ellipse can somewhat mitigate this issue, it cannot eliminate the interference from mining areas, creating notable challenges for subsequent machine learning classifications.

In past study, most scholars have relied on prior knowledge of fire zone ranges, focusing on smaller study areas. For broader studies, visual interpretations have been used to rule out mining areas, but this introduces a level of subjectivity. Consequently, a deeper investigation into distinguishing the formation mechanisms of fire areas from mining areas is a pivotal direction for future study.

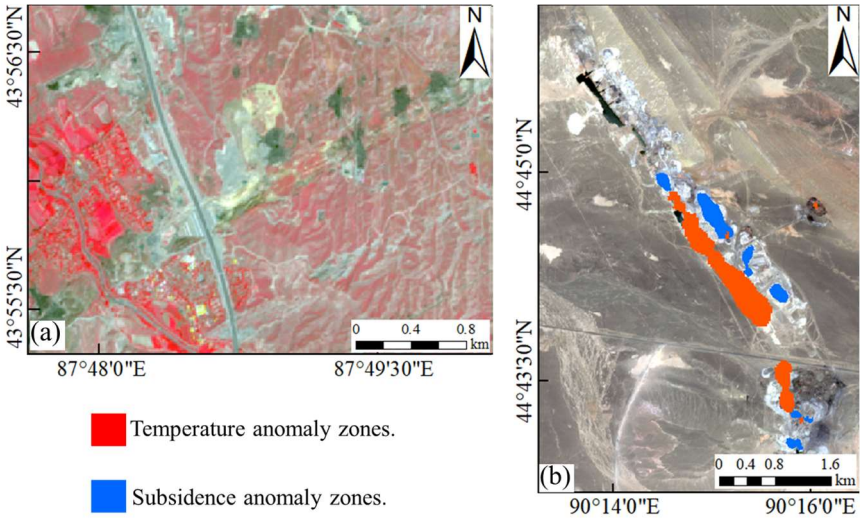

Fig. 15. The extent of the anomaly overlapping. (a) The final overlap of temperature and subsidence anomalies in the Sandaoba fire area. (b) The final overlap of temperature and subsidence anomalies in the Jiangjun Gobi and the northern fire areas of Jiangjun Temple (there is no overlap in the southern fire area of Jiangjun Temple).

B. The Limitations of Multi-source Remote Sensing in Fire Zone Identification Methods

Utilizing multi-source remote sensing information, various anomalies can be detected across multiple dimensions without direct contact with the target. Through collaborative analysis of information and related data processing, it is possible to identify the region of fire zones. However, there are still several challenges:

1) Temporal Discrepancies: Multi-source remote sensing images are widely dispersed in time. For instance, the night light data referenced in this study is available for only one month in 2018. This temporal limitation restricts the detection capabilities of open fire areas and complicates the task of differentiating interference areas based on brightness levels.

2) Atmospheric Interference: Multispectral images can be compromised by atmospheric conditions like clouds and fog, leading in poor imaging quality for certain periods.

3) Resolution Issues: Image resolution plays a critical role in the fine identification of the fire zones. This study employed four types of remote sensing data, the finest

resolution of which is just 10m. As shown in Table V, the western fire area of Sandaoba is less than 10,000m^2, translating to a size of 10×10 pixels in imagery. With a lower resolution, anomaly and background information tend to merge within a single pixel, making it challenging to distinguish between the anomalous areas and their boundaries. Such errors, originating at the initial stage, can escalate the risk of introducing non-fire zone samples when delineating fire zone samples using the standard deviation ellipse. This can mislead the classifier into erroneous learning, ultimately resulting in misjudgments and missed detections when identifying fire zones. Therefore, while trying to use high-resolution images as much as possible, the method of sample delineation should also consider both the quality and quantity of the results.

4) Efficiency and Accuracy of the Model: Although the AdaBoost_RF_StBP model constructed in this study does enhance classification accuracy compared to individual classifiers. However, having more classifiers will increase processing time. Therefore, finding a balance between the number of classifiers and their accuracy is worth further exploration.

5) The Monitoring of New Fire Focus: For the same fire area, by periodically adding data from areas of interest, we can expand the number of training samples and achieve continuous monitoring of long time series. However, we are more concerned that the added new samples should be efficiently trained with the old samples, rather than starting training from scratch, in order to save time.

By addressing these points, researchers can aim for more accurate and efficient remote sensing methods for fire area identification.

VI. Conclusions

This study explores multi-modal remote sensing data for the multi-factor extraction of fire zones, delineates features of concealed fire zones, open fire zones, and non-fire zones, and constructs a AdaBoost_RF_StBP coal fire zone identification model. These three elements form the core of the study content, providing a novel approach for the identification of coal fire zones.

1) Using a diverse array of data - thermal infrared, multispectral, SAR, and night light - we extracted temperature anomaly zones, vegetation-rich zones, deformation anomaly zones, and high-brightness zones. These are extracted by adding or subtracting n times the standard deviation to the mean at varying confidence levels. This strategy curtails the reliance on pre-existing knowledge of fire zones and reduces the subjectivity of threshold determination.

2) The fire zones in Jiangjun Gobi and Jiangjun Temple have experienced relatively brief burning durations. Consequently, the overlap areas between temperature anomalies and subsidence anomalies are limited. By constructing a 1 times standard deviation ellipse in the intersecting areas of the two, it is possible to assist in delineating the range of concealed fire zone samples and increase the sample quantity.

3) Night light data reveals stark brightness discrepancies between regions. The maximum brightness in the Jiangjun Gobi and Jiangjun Temple fire zones peaks at 0.03920, in contrast to the Beishan coal mines, which caps at a mere 0.005320. The significant brightness difference between fire zones and mining areas indicates that night light data has great potential for distinguishing between them.

4) The AdaBoost_RF_StBP coal fire zone identification model enhances traditional algorithms by integrating Bagging, Boosting, and Stacking concepts. This mitigates bias, variance, and overfitting in predictions. The initial base-learner has an accuracy between 0.56 and 0.87, but the meta-learner boosts this to 0.8-0.96. From the final identification results, the recognition areas of AdaBoost_RF_StBP in Jiangjun Gobi, Jiangjun Temple and Sandaoba fire areas account for 122.7%, 109.8% and 115.8% of the actual fire areas, respectively. The redundant areas are mostly excavation pits proximate to the fire zones and could potentially mark the zones of fire spread. As a validation area for model generalization, the Sandaoba fire area has a misjudgment rate of less than 20%, demonstrating the model possesses high applicability to different fire zones. Overall, combining multi-source remote sensing data with the AdaBoost_RF_StBP coal fire zone identification model, this study introduces an effective means for fire zone identification and monitoring.

Acknowledgement

The authors would like to thank the European Space Agency (ESA) for their generosity in providing the Sentinel-1A SAR data and Sentinel-2A Multispectral data, the United States Geological Survey (USGS) for their generosity in providing the Landsat-8 Thermal infrared data, the Luo Jia Laboratory for their generosity in providing the Luojia-01 data, Yu et al. for providing GACOS for free. We appreciate the help from postgraduates Jie Li, Si Chen, and Weiyun Cao.

References

[1] G. B. Stracher, A. Prakash, and G. Rein, "Chapter 1 - Smoldering-Peat Megafires: The Largest Fires on Earth," in Coal and Peat Fires: a Global Perspective," *Eds. Boston: Elsevier*, pp. 1-11, 2015.

[2] G. B. Stracher and T. P. Taylor, "Coal fires burning out of control around the world: thermodynamic recipe for environmental catastrophe," *International Journal of Coal Geology*, vol. 59, no. 1, pp. 7-17, Jul. 2004.

[3] K. S. Brown, "Subterranean Coal Fires Spark Disaster," Science, vol. 299, pp. 1177 - 1177, 2003.

[4] C. Kuenzer and G. B. Stracher, "Geomorphology of coal seam fires," *Geomorphology*, vol. 138, no. 1, pp. 209-222, Feb. 2012.

[5] H. Guan and J. V. Genderen et al., "Investigation and Research on Spontaneous Combustion Environment in Coalfields of Northern China,". *Investigation and Research on Spontaneous Combustion Environment in Coalfields of Northern China*, 1998.

[6] M. A. Engle et al., "Gas emissions, minerals, and tars associated with three coal fires, Powder River Basin, USA," *Science of The Total Environment*, vol. 420, pp. 146-159, Mar. 2012.

[7] S. M. M. Mahfuz et al., "Risk assessment of naturally occurring radioactivity in soil adjacent to a coal-fired brick kiln," *Radiation Physics and Chemistry*, vol. 209, 2023.

[8] Y. Wang et al., "Research Progress and Prospect on Ecological Disturbance Monitoring in Mining Area," *Acta Geodaetica et Cartographica Sinica*, vol. 46, no. 10, pp. 1705-1716, 2017.

[9] J. Deng, S. Ge, H. Qi, F. Zhou, and B. Shi, "Underground coal fire emission of spontaneous combustion, Sandaoba coalfield in Xinjiang, China: Investigation and analysis," *Science of The Total Environment*, vol. 777, p. 146080, Jul. 2021.

[10] X. Chen, J. Peng, Z. Song, Y. Zheng, and B. Zhang, "Monitoring Persistent Coal Fire Using Landsat Time Series Data From 1986 to 2020," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16, 2022.

[11] M. Karaoulis, A. Revil, and D. Mao, "Localization of a coal seam fire using combined self-potential and resistivity data," International Journal of Coal Geology, vol. 128-129, pp. 109-118, Aug. 2014.

[12] B. Kong, E. Wang, W. Lu, and Z. Li, "Application of electromagnetic radiation detection in high-temperature anomalous areas experiencing coalfield fires," Energy, vol. 189, p. 116144, Dec. 2019.

[13] Z. Lu et al., "Study on the formation mechanism of a radon source

during coal spontaneous combustion in a goaf," Fuel, vol. 336, p. 127135, Mar. 2023.

[14] F. Cao, M. Meng, B. Shan, and R. Sun, "Source apportionment of mercury in surface soils near the Wuda coal fire area in Inner Mongolia, China," Chemosphere, vol. 263, p. 128348, 2021.

[15] R. Sun et al., "Atmospheric Mercury Isotope Shifts in Response to Mercury Emissions from Underground Coal Fires," Environmental Science & Technology, vol. 57, no. 23, pp. 8638-8649, Jun. 2023.

[16] R. S. Chatterjee, "Coal fire mapping from satellite thermal IR data – A case example in Jharia Coalfield, Jharkhand, India," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 113-128, Apr. 2006.

[17] Z. Song et al., "Analysis of coal fire dynamics in the Wuda syncline impacted by fire-fighting activities based on in-situ observations and Landsat-8 remote sensing data," *International Journal of Coal Geology*, vol. 141-142, pp. 91-102, Mar. 2015.

[18] Z. Yang et al., "Locating and defining underground goaf caused by coal mining from space-borne SAR interferometry," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 135, pp. 112-126, Jan. 2018.

[19] L. Zhou, D. Zhang, J. Wang, Z. Huang, and D. Pan, "Mapping Land Subsidence Related to Underground Coal Fires in the Wuda Coalfield (Northern China) Using a Small Stack of ALOS PALSAR Differential Interferograms," *Remote Sensing*, vol. 5, issue 3, pp. 1152-1176, vol. 5, pp. 1152-1176, Mar. 2013.

[20] Y. Chen et al., "Revealing Land Surface Deformation Over the Yineng Backfilling Mining Area, China, by Integrating Distributed Scatterer SAR Interferometry and a Mining Subsidence Model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3611-3634, 2023.

[21] R. Ghosh, P. K. Gupta, V. A. Tolpekin, S. K. Srivastav, and S. Majumdar, "The Potential of Channel Specific Reflectance in Landsat 8 OLI Sensor for Retrieving Coal Fire Affected Pixels," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 9968-9971, 2019.

[22] R. Ghosh, P. K. Gupta, V. Tolpekin, and S. K. Srivastav, "An enhanced spatiotemporal fusion method – Implications for coal fire monitoring using satellite imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 88, p. 102056, Jun. 2020.

[23] V. Karanam, M. Motagh, S. Garg, and K. Jain, "Multi-sensor remote sensing analysis of coal fire induced land subsidence in Jharia Coalfields, Jharkhand, India," *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, p. 102439, Oct. 2021.

[24] J. Liu et al., "Underground Coal Fire Detection and Monitoring Based on Landsat-8 and Sentinel-1 Data Sets in Miquan Fire Area, XinJiang," *Remote Sensing*, vol. 13, p. 1141, Mar. 2021.

[25] T. T. Nguyen, T. D. Vu, and H. V. Environment, "Use of Hot Spot Analysis to Detect Underground Coal Fires from Landsat-8 TIRS Data: A Case Study in the Khanh Hoa Coal Field, North-East of Vietnam," *The Environment and Natural Resources Journal*, 2019.

[26] L. Jiang, H. Lin, J. Ma, B. Kong, and Y. Wang, "Potential of small-baseline SAR interferometry for monitoring land subsidence related to underground coal fires: Wuda (Northern China) case study," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 257-268, Feb. 2011.

[27] R. M. Jameela, S. T. Hassan, K. Hrishikesh, and K. Claudia, "Detecting and Analyzing the Evolution of Subsidence Due to Coal Fires in Jharia Coalfield, India Using Sentinel-1 SAR Data," *Remote Sensing*, vol. 13, no. 8, 2021.

[28] Y. Wang et al., "Research on Multi-source Remote Sensing Detection of Concealed Fire Sources in Coalfields," *Geomatics and Information Science of Wuhan University*, vol. 47, no. 10, pp. 1651-1661, 2022.

[29] B. Yu et al., "Coal fire identification and state assessment by integrating multitemporal thermal infrared and InSAR remote sensing data: A case study of Midong District, Urumqi, China," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 190, pp. 144-164, Aug. 2022.

[30] T. Wang et al., "A spatio-temporal temperature-based thresholding algorithm for underground coal fire detection with satellite thermal infrared and radar remote sensing," *International Journal of Applied Earth Observation and Geoinformation*, vol. 110, p.102805, Jun. 2022.

[31] Y. Chen et al., "Active–Passive Remote Sensing Identification of Underground Coal Fire Zones With Joint Constraints of Temperature and Surface Deformation Time Series," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 894-915, 2024.

[32] F. Zhao and J. J. Mallorqui, "Coherency Matrix Decomposition-Based Polarimetric Persistent Scatterer Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7819-7831, 2019.

[33] C. Zhang and J. Zhang, "A Survey of Selective Ensemble Learning Algorithms," *Chinese Journal of Computers*, vol. 34, no. 08, pp. 1399-1410, 2011.

[34] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 166-177, Jun. 2019.

[35] G. M.A., H. Minghui, M. A.K., T. M., and S. P.N., "Ensemble deep learning: A review," *Engineering Applications of Artificial Intelligence*, vol. 115, 2022.

[36] Y. Zhang, S. Liang, Z. Zhu, H. Ma, and T. He, "Soil moisture content retrieval from Landsat 8 data using ensemble learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 185, pp. 32-47, Mar. 2022.

[37] Z. Qin, Z. Minghua, A. Karnieli, and P. Berliner, "Mono-window Algorithm for Retrieving Land Surface Temperature from Landsat TM6 data," *Acta Geographica Sinica*, no. 04, pp. 456-466, 2001.

[38] Q. Vanhellemont, "Combined land surface emissivity and temperature estimation from Landsat 8 OLI and TIRS," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 166, pp. 390-402, Aug. 2020.

[39] R. Niclòs et al., "Evaluation of Landsat-8 TIRS data recalibrations and land surface temperature split-window algorithms over a homogeneous crop area with different phenological land covers," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 174, pp. 237-253, Apr. 2021.

[40] L. j et al., "Underground Coal Fires Identification and Monitoring Using Time-Series InSAR With Persistent and Distributed Scatterers: A Case Study of Miquan Coal Fire Zone in Xinjiang, China," *IEEE Access*, vol. 7, pp. 164492-164506, 2019.

[41] B. P., F. G., L. R., and S. E., "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, 2002.

[42] B. Osmanoğlu, F. Sunar, S. Wdowinski, and E. Cabral-Cano, "Time series analysis of InSAR data: Methods and trends," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 90-102, May. 2016.

[43] N. Yague-Martinez et al., "Interferometric Processing of Sentinel-1 TOPS Data, *IEEE T. Geosci. Remote*, 54, 2220–2234," ed, 2016.

[44] R. M. Goldstein and H. A. Zebker, "Interferometric radar measurement of ocean surface currents," *Nature*, vol. 328, no. 6132, pp. 707-709, Aug 1987.

[45] N. Levin, K. Johansen, J. M. Hacker, and S. Phinn, "A new source for high spatial resolution night time images — The EROS-B commercial satellite," *Remote Sensing of Environment*, vol. 149, 2014.

[46] Z. Chen et al., "A New Approach for Detecting Urban Centers and Their Spatial Structure With Nighttime Light Remote Sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, pp. 1-15, Aug. 2017.

[47] B. Wu, C. Yang, Q. Wu, C. Wang, J. Wu, and B. Yu, "A building volume adjusted nighttime light index for characterizing the relationship between urban population and nighttime light intensity," *Computers, Environment and Urban Systems*, vol. 99, p. 101911, Jan. 2023.

[48] L. Han, L. Wang, X. Ding, H. Wen, X. Yuan, and W. Zhang, "Similarity quantification of soil parametric data and sites using confidence ellipses," *Geoscience Frontiers*, vol. 13, no. 1, p. 101280, Jan. 2022.

[49] T. Zeng, L. Wu, D. Peduto, T. Glade, Y. S. Hayakawa, and K. Yin, "Ensemble learning framework for landslide susceptibility mapping: Different basic classifier and ensemble strategy," *Geoscience Frontiers*, vol. 14, no. 6, p. 101645, Nov. 2023.

[50] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," Engineering Applications of Artificial Intelligence, vol. 115, p. 105151, Oct. 2022.

[51] J. L. Davis, T. A. Herring, I. I. Shapiro, A. E. E. Rogers, and G. Elgered, "Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length," (in English), *Radio Science, Article* vol. 20, no. 6, pp. 1593-1607-1607, Jan. 1985.

Yu Chen received the master degree in Photogrammetry and Remote Sensing from China University of Mining and Technology (CUMT), Xuzhou, China, in 2012, and Ph.D. degree in earth and planetary science from University of Toulouse, Toulouse, France, in 2017.

She worked as an assistant researcher in the GET (Géoscience Environnement Toulouse) laboratory of the French National Center for Scientific Research in 2013. She is currently an associate professor in CUMT. Her research interest is remote sensing information extraction and application, SAR interferometry, with particular emphasis on its application for geophysical studies.

Kaiwen Ding received the B.E. degree in surveying and mapping engineering from the Shandong University of Science and Technology, Tai'an, China, in 2022. And he is a graduate student majoring in surveying and mapping engineering at the China University of Mining and Technology, Xuzhou, China.

His research interests are remote sensing image processing and machine learning data analysis.

Zhihui Suo received the B.E. degree in surveying and mapping engineering from the Henan Polytechnic University, Jiaozuo, China, in 2021. And he is also a graduate student majoring in surveying and mapping engineering with the China University of Mining and Technology, Xuzhou, China.

His research interests are identification of underground coal fires and inversion of cavity parameters.

Kun Tan (SM'16**)** received the B.S. degree in information and computer science from Hunan Normal University, Hunan, China, in 2004, and the Ph.D. degree in photogrammetric and remote sensing from China University of Mining and Technology, Jiangsu, China, in 2010. He was a Joint Ph.D. candidate in Remote Sensing at Columbia University, USA during September 2008 to September 2009.

From 2010 to 2018, he has been with the Department of Surveying, Mapping and Geoinformation, China University of Mining and Technology, Xuzhou City, Jiangsu Province, China. He is currently a Professor in East China Normal University. His research interests include hyperspectral image classification and detection, spectral unmixing, quantitative inversion of land surface parameters, and urban remote sensing.

Jun Wei received the B.E. degree in surveying and mapping engineering from the Xi'an University of Science and Technology, Xian, China, in 1995. He is a senior engineer at the Xinjiang Uygur Autonomous Region Mine Safety Service and Guarantee Center, Urumqi 830017, China.

His research interest is technical management of mine safety.

Fei Cao, received the B.E. degree in surveying and mapping engineering from the Chengdu University of Technology, Chengdu, China. And he is a senior engineer at the Xinjiang Uygur Autonomous Region Mine Safety Service and Guarantee Center, Urumqi 830017, China.

His research interests include surveying and preventing underground coal fires.

Huahai Sun received the B.E. degree in resource exploration engineering from the Xinjiang University, Urumqi, China, in 2009. And he is a senior engineer at the Xinjiang Uygur Autonomous Region Mine Safety Service and Guarantee Center, Urumqi 830017, China.

His research interests include prevention and early warning of coalfield fires.

Mengmeng Bu received the B.E. degree in geographical information science from the the China University of Mining and Technology, Xuzhou, China, in 2010, and the MA.Eng from the China University of Mining and Technology, Xuzhou, China, in 2020. She is a senior engineer at the Fifth Geological Brigade of the Bureau of Geology and Mining, Xuzhou 221004, China.

Her research interest is geotechnical engineering geohazard prevention and control.