Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 7 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i07.48601

PAPER

# Advancements in Satellite Remote Sensing for Predicting Large-Scale Wildfire Risks: An Image Processing Algorithmic Framework

**Boxin Li[1], Hong'e Ren[1](✉), Jing Tian[2]**

[1]College of Computer and Control Engineering, Northeast Forestry University, Harbin, China

[2]College of Surveying and Mapping Engineering, Heilongjiang Institute of Technology, Harbin, China

rhe@nefu.edu.cn

**ABSTRACT**

With the escalation of global warming and human activities, large-scale wildfires have become increasingly frequent, posing significant threats to both ecological environments and human societal safety. Satellite remote sensing technology plays a pivotal role in wildfire monitoring and risk assessment, providing extensive geographical coverage and continuous monitoring capabilities. Traditional methods for predicating wildfire risk, however, face limitations in processing large-scale remote sensing data, especially in cloud detection and temporal information analysis. In response to this challenge, a novel set of image processing algorithms has been developed to enhance the efficiency and accuracy of wildfire risk prediction. Initially, a cloud detection and removal algorithm based on deep learning is introduced. This algorithm effectively identifies and eliminates cloud interference in remote sensing images, thereby significantly improving the quality and usability of image data. Subsequently, a temporal information capturing technique is proposed, capable of processing vast amounts of remote sensing data and extracting time series features. This technique provides robust data support for wildfire risk prediction. The application of these technologies not only improves the efficiency of the data processing workflow but also enhances the timeliness and accuracy of the prediction model, holding significant practical importance for guiding actual wildfire prevention and response measures.

**KEYWORDS**

wildfire risk prediction, satellite remote sensing, image processing algorithm, cloud detection, temporal information, deep learning

## 1 INTRODUCTION

Under the dual impact of global climate change and human activities, the frequency and intensity of extreme wildfire events have significantly increased, posing severe threats to the natural environment and socio-economic stability [1, 2]. Satellite

Li, B., Ren, H., Tian, J. (2024). Advancements in Satellite Remote Sensing for Predicting Large-Scale Wildfire Risks: An Image Processing Algorithmic Framework. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(7), pp. 19–33. https://doi.org/10.3991/ijim.v18i07.48601

Article submitted 2023-12-08. Revision uploaded 2024-1-28. Final acceptance 2024-02-13.

remote sensing technology, with its extensive monitoring range and high-frequency data acquisition capability, plays an irreplaceable role in large-scale wildfire risk assessment and prediction [3, 4]. Through the analysis and processing of remote sensing images, vegetation conditions and environmental changes can be monitored in a timely manner. This process effectively identifies areas at risk of wildfires, providing a scientific basis for making decisions related to disaster prevention and mitigation [5, 6].

However, the complexity of wildfire risk prediction arises not only from the uncertainty of wildfires themselves but also from the extensive amount of remote sensing data that needs to be processed. This data is sourced from multiple channels at various scales and is susceptible to interference factors such as cloud cover [7]. Traditional methods of wildfire risk assessment are often limited to specific regions and struggle to adapt to the requirements of large-scale, dynamic risk prediction [8–10]. Therefore, the development of a new image processing algorithm capable of efficiently utilizing satellite remote sensing data for large-scale wildfire risk prediction is an urgent need in the field of natural disaster monitoring [11].

Existing research methods in remote sensing image processing for wildfire risk prediction have several deficiencies. Cloud cover significantly impacts the usability of remote sensing images. Existing cloud detection and removal techniques lack the accuracy and real-time performance necessary for effective fire warning [12–14]. Secondly, the capture of temporal information often overlooks the continuity and dynamic changes in the time dimension, leading to predictions that lack timeliness and accuracy [15–17]. In light of these issues, the development of more efficient and intelligent image processing methods is imperative.

This study introduces a cloud detection and removal algorithm for satellite remote sensing images, targeting large-scale wildfire risk prediction. Utilizing deep learning technology, this algorithm can accurately identify clouds against complex backgrounds and effectively remove their influence from remote sensing data, thereby enhancing the usability of images. Secondly, this study introduces a novel technique for capturing temporal information. This technique is capable of processing large volumes of remote sensing data and extracting key time series features to construct more accurate wildfire risk prediction models. This study optimizes the data preprocessing workflow. It improves the timeliness and accuracy of wildfire risk predictions. Additionally, it establishes a new technical approach for remote sensing and monitoring of similar natural disasters. This study holds significant theoretical and practical value.

## 2 CLOUD DETECTION AND REMOVAL OF SATELLITE REMOTE SENSING IMAGES FOR LARGE-SCALE WILDFIRE RISK PREDICTION

In light of the frequent occurrence of large-scale wildfires worldwide, the utilization of satellite remote sensing data for risk prediction becomes especially critical. However, cloud cover often affects remote sensing images, severely restricting the accuracy and timeliness of wildfire monitoring. Figure 1 illustrates the workflow for large-scale wildfire risk prediction based on satellite remote sensing data. We understand that researching cloud detection and removal in remote sensing images for large-scale wildfire risk prediction significantly improves the quality and usability of image data. This provides a clearer and

more reliable visual basis for subsequent fire analysis. Additionally, it strengthens the response speed and prediction accuracy of fire warning systems, offering strong technical support to mitigate the ecological and socio-economic losses caused by wildfires.

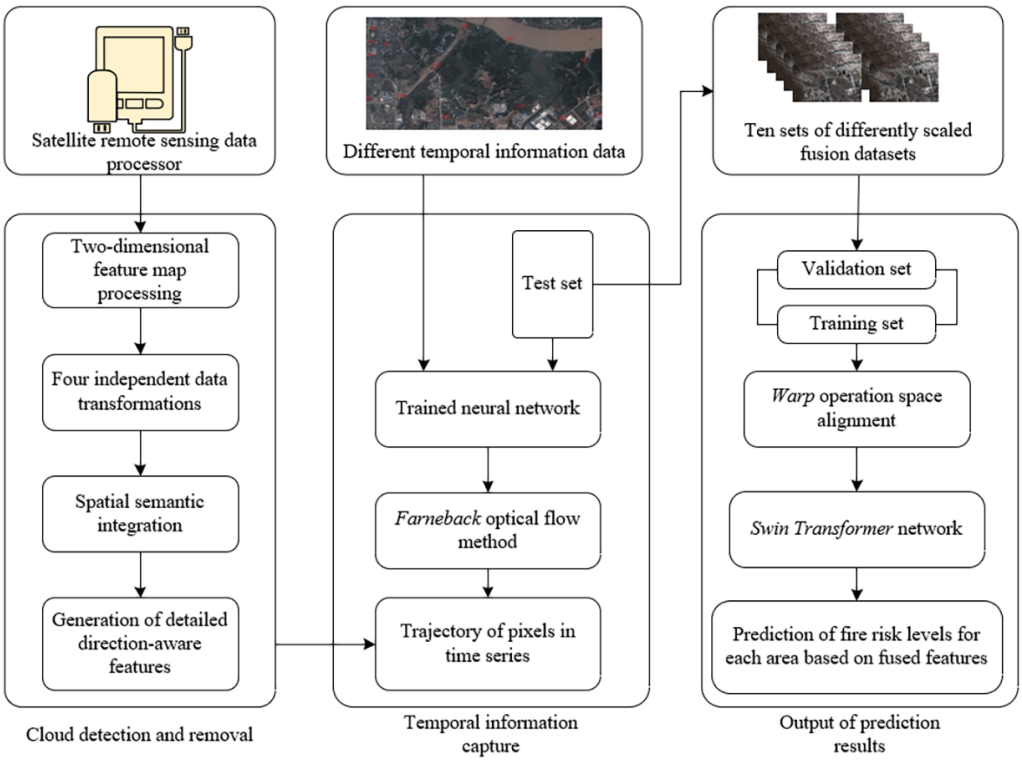


**Fig. 1.** Workflow for large-scale wildfire risk prediction based on satellite remote sensing data

## 2.1 Cloud detection method

A cloud detection method for satellite remote sensing images aimed at predicting large-scale wildfire risk is proposed in this study. The basic principle involves extending the concept of traditional recurrent neural networks (RNN) into the two-dimensional spatial domain of image data processing to better handle and analyze data with spatial correlations. Figure 2 presents a schematic diagram of the cloud detection network structure.

The method initially processes the input two-dimensional feature map with a $1 \times 1$ convolution kernel. The purpose of this step is to transform the original input features into a data format suitable for the hidden layer, preparing for subsequent spatial correlation analysis. This transformation reduces data dimensions while preserving essential features, establishing an appropriate data foundation for further processing. Subsequently, the model applies four independent data transformations corresponding to the four main directions of an image (left, right, up, and down). These transformations operate independently, aggregating spatial semantic information along their respective directions. The critical step lies in capturing the spatial dependencies of the image in various directions, thereby effectively extracting structural information from the image, such as the boundaries and shapes of cloud layers.

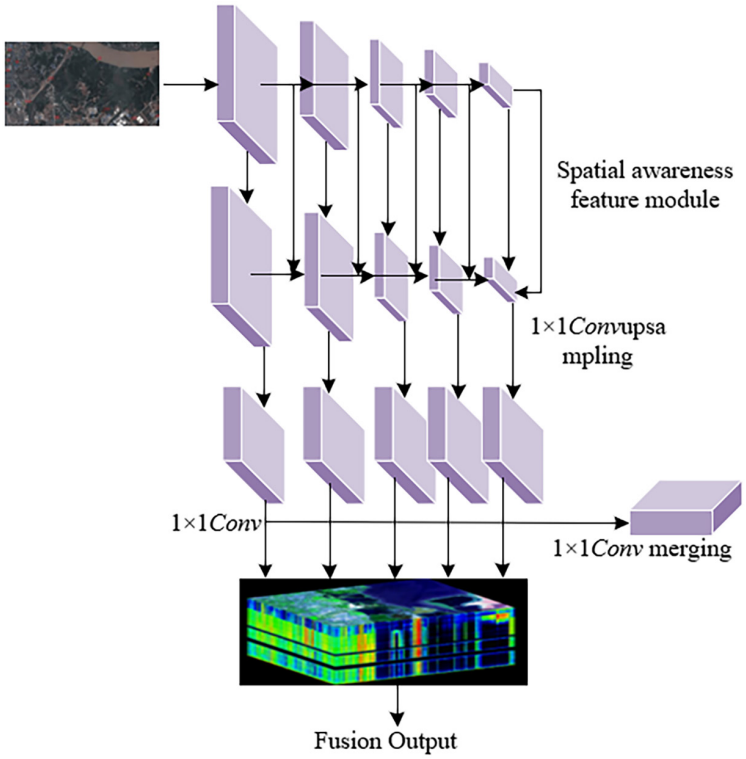


**Fig. 2.** Schematic diagram of cloud detection network structure

To execute data transformations within the spatial RNN, the iterative neural network (INN) model is introduced in this section. It is assumed that the feature of pixel ($u$, $k$) is represented by $g_{uk}$, with the feature map width and the weight parameters for cyclic transformation towards a specific direction denoted by $v$ and $x_{RI}$, respectively. A round of data transformation in a specific direction can be accomplished by repeatedly performing the operation illustrated in equation (1).

$$g_{u.k} = MAX(\beta_{RI} g_{u,k-1} + g_{u,k}, 0) \tag{1}$$

Furthermore, the spatial semantics gathered from these four directions are consolidated into an intermediate feature map. This fusion process combines information from various directions to create a more comprehensive representation of spatial semantics. The aforementioned process is repeated to further propagate and aggregate spatial contexts in each principal direction. Through iteration, the model enhances the spatial relationships within the feature map, making the spatial context information more distinct and improving the model's ability to detect features such as cloud layers.

We suggest using a direction-aware attention mechanism. This will help improve cloud detection accuracy in satellite remote sensing images. It will also enhance the wildfire risk prediction model's understanding of spatial environmental characteristics. The structure of this mechanism is depicted in Figure 3. The core idea of the direction-aware attention mechanism lies in recognizing that features in different spatial directions have varying degrees of importance for cloud identification and wildfire monitoring. To capture spatial features across different directions, the mechanism selectively emphasizes or suppresses features in specific directions through attention weights learned by the spatial RNN. Thus, the model is enabled to focus more on information crucial for the specific task, thereby improving the accuracy of cloud detection and wildfire risk prediction.

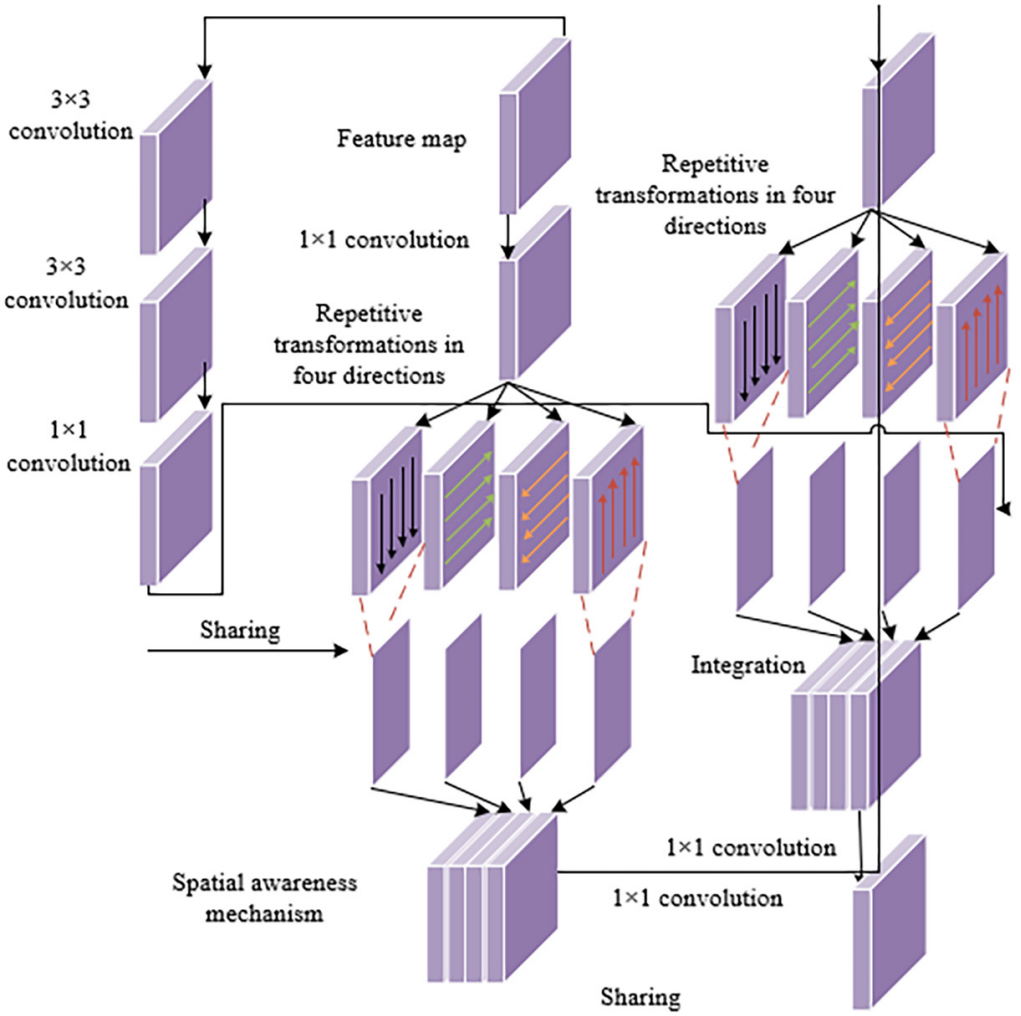


**Fig. 3.** Schematic diagram of the direction-aware attention mechanism module

The model initially employs two consecutive convolution layers followed by ReLU non-linear operations to extract spatial features. An attention weight is then generated through a third convolution layer. The generated attention weights are divided into four weight maps corresponding to the four principal spatial directions (left, right, up, down), denoted as $Q_{LE}$, $Q_{DO}$, $Q_{RI}$, and $Q_{UP}$. Letting the attention estimation network operator be represented by $d_{AT}$, the input feature mapping by $a$, and the learnable parameters by $\varphi$, the following is obtained:

$$q = d_{AT}(a;\varphi) \tag{2}$$

These weight maps undergo element-wise multiplication with spatial environmental features, enabling the model to selectively utilize these features as it moves in different directions. The attention-weighted features are then integrated and passed through a 1×1 convolution to reduce the feature dimensions to a quarter of the original, size facilitating the simulation of the transformation of hidden data. A second round of cyclic transformation is executed, using the same attention weights previously learned to further select and emphasize spatial semantics. Finally, a 1×1 convolution and ReLU non-linear operation are applied to the integrated features, which are mapping to simulate the transformation from hidden layers

to output, generating the final direction-aware spatial semantic features. Through such processing, the model is capable of generating detailed direction-aware features, crucial for distinguishing between cloud and non-cloud features. This in turn, provides more accurate input data for wildfire risk prediction.

Regarding the loss function, a combination of loss functions is employed at every layer of the cloud detection network to balance the detection of cloud and non-cloud pixels. This approach helps the network focus on learning classes that are difficult to classify. Assuming the weights and loss of the entire network at layer $u$ are represented by $q_u$ and $M_u$, and the weights and loss of the multi-layer composite characteristic layers by $q_l$ and $M_l$, with the fusion layer's weights and loss represented by $q_d$ and $M_d$, the expression for the overall loss function is given as follows:

$$M_{OA} = \sum_{u} q_u M_u + q_l M_l + q_d M_d \tag{3}$$

## 2.2 Cloud removal method

Traditional cloud detection algorithms may require cloud masks as training supervision signals. In this work, the choice was made to use clear, cloud-free images as the supervision signal. This approach enables the network to learn how to recover cloud-free images from those containing clouds. Cloud-free images directly guide the network in removing cloud pixels and restoring the information obscured by the clouds. This approach aims for the network to generate results very close to real cloud-free images, rather than merely identifying cloud locations. In image reconstruction tasks, the Euclidean loss function is more commonly used than the weighted cross-entropy loss function. Hence, the Euclidean loss function is adopted to optimize the network. This loss function directly measures the pixel-level differences between the predicted image and the target cloud-free image, thus focusing the model on generating visually accurate cloud-free images. Additionally, an optimizer based on the Markov decision process (MDP) for neural networks is constructed. Such an optimizer can be used to dynamically adjust parameters during the network training process, allowing the network to more effectively learn how to remove clouds, particularly in the presence of complex terrestrial and cloud boundary conditions. By combining these two aspects, the proposed cloud removal method can more accurately handle cloud interference in remote sensing images by directly learning to reconstruct cloud-free images. This provides clearer input data for wildfire risk prediction, increasing the accuracy and reliability of the model's fire predictions.

Clouds affect the color consistency of remote sensing images. This leads to inaccuracies in extracting image features, which affects the prediction outcomes of wildfire risk models. It's crucial to remove visual interference caused by clouds. This is crucial for the wildfire risk prediction model to accurately interpret and analyze remote sensing images. Therefore, a compensation mechanism is introduced to correct the color bias caused by clouds using a color transfer function, ensuring the quality and consistency of the model input data. Thus, the model can focus more on the actual conditions of the Earth's surface, such as vegetation status and surface temperature, which are key indicators. This enhances the accuracy and reliability of wildfire risk predictions. Assuming the cloud images and cloud-free images used for training are represented by $U_t$ and $U_v$, respectively, and the cloud and non-cloud regions in the image space by $\Psi_t$ and $\Psi_v$, the color transfer function by $S_d$, and the

color compensation error by $R_z$. The goal is to find $S_d$ that minimizes $R_z$ between $\Psi_t$ and $\Psi_v$ on $U_v$.

$$R_z = |U_t - S_d(U_v)|^2_{\Psi_v} \tag{4}$$

Assuming the color value of a pixel is denoted by $a$, and the parameters in the color transfer function are characterized by $(e,h,y)$ and $L_\beta$, then the linear expression for $S_d$ can be represented as:

$$S_d(a) = L_\beta \cdot \begin{pmatrix} e \\ h \\ y \\ 1 \end{pmatrix} \tag{5}$$

Specifically, regarding the loss function, assuming the network prediction is $U_v$, the loss across the entire image domain, $M^e$, is calculated through the color space of $S_d(U_v)$ and $U_v$:

$$M^e = |S_d(U_v - \tilde{U}_v)|^2_{\Psi_v \cup \Psi_t} \tag{6}$$

In every layer of the cloud removal network, the loss function calculated through the color space of $S_d(U_v)$ and $U_v$ is employed, resulting in the expression of the total loss function $M^e_{OA}$:

$$M^e_{OA} = \sum_u q^e_u L^e_u + q^e_l M^e_l + q^e_d M^e_{ld} \tag{7}$$

## 3 CAPTURING TEMPORAL INFORMATION IN SATELLITE REMOTE SENSING IMAGES FOR LARGE-SCALE WILDFIRE RISK PREDICTION

After detecting and removing clouds in satellite remote sensing images, it is essential to explore capturing temporal information for large-scale wildfire risk prediction. This is attributed to the fact that fire risk assessments rely not only on the surface conditions at a single moment but also require an understanding of the surface's changing trends over time. This includes vegetation growth cycles, seasonal variations, drought levels, and the long-term impacts of other environmental factors. Through comprehensive analysis of temporal information, it is possible to predict fire occurrences more accurately. This can provide a more scientific foundation for early warning systems, resource distribution, and disaster response strategies. Moreover, temporal analysis helps identify and interpret long-term trends and sudden events that may elevate fire risk, thereby improving the effectiveness of fire prevention and response measures. The specific steps involved in capturing temporal information in satellite remote sensing images for large-scale wildfire risk prediction are detailed below, as illustrated in Figure 4.

Step 1: Initially, the neural network receives and processes consecutive moments of satellite remote sensing image data. Before processing the current moment, data from previous moments had already passed through the neural network. Each layer extracts feature maps representing image content, such as surface temperature, vegetation status, and other factors associated with fire risk. These feature maps play a crucial role in subsequent steps, as they contain key information about the surface

conditions from the previous moment. This information is indispensable for monitoring environmental changes and predicting fire risk.

Step 2: The utilization of the Farneback optical flow method to estimate pixel movement between two consecutive frames is essential for comprehending dynamic changes on the surface. Optical flow provides the trajectory of pixel movement in a time series, which is crucial for capturing subtle surface changes, such as the growth or wilting of vegetation and changes in ground moisture. These changes may indicate an increase in the risk of fire. Assuming the gradients of the image in directions $a$ and $b$ are represented by $U_a(a,b)$ and $U_b(a,b)$, and the difference in grayscale values between two adjacent frames by $U_s(a,b)$. The time interval between two adjacent frames and the scale factor between different levels in the image pyramid are denoted by $\Delta_s$ and $\Delta_a$. Based on the following equations, the optical flow vector for each pixel in the current pyramid level of the satellite remote sensing image can be obtained:

$$U_a(a,b) = \frac{1}{4\Delta_s} U \begin{pmatrix} I(a+1,b,s_1) + U(a+1,b,s_2) \\ -U(a-1,b,s_1) - U(a-1,b,s_2) \end{pmatrix} \tag{8}$$

$$U_b(a,b) = \frac{1}{4\Delta_s} \begin{pmatrix} U(a,b+1,s_1) + U(a,b+1,s_2) \\ -U(a,b-1,s_1) - U(a,b-1,s_2) \end{pmatrix} \tag{9}$$

$$U_s(a,b) = \frac{1}{4\Delta_s} \begin{pmatrix} U(a+1,b,s_2) - U(a-1,b,s_2) \\ +U(a+1,b,s_1) - U(a-1,b,s_1) \end{pmatrix} \tag{10}$$

Step 3: We use motion information obtained from optical flow calculations. We align feature maps from the previous time step with current satellite remote sensing images through a warp operation. This step ensures that surface features observed between consecutive frames of satellite remote sensing images can be accurately compared and analyzed. Spatio-temporal alignment is important for understanding temporal changes. It removes biases caused by changes in satellite observation angles or natural surface movements. It keeps only the differences induced by environmental changes.

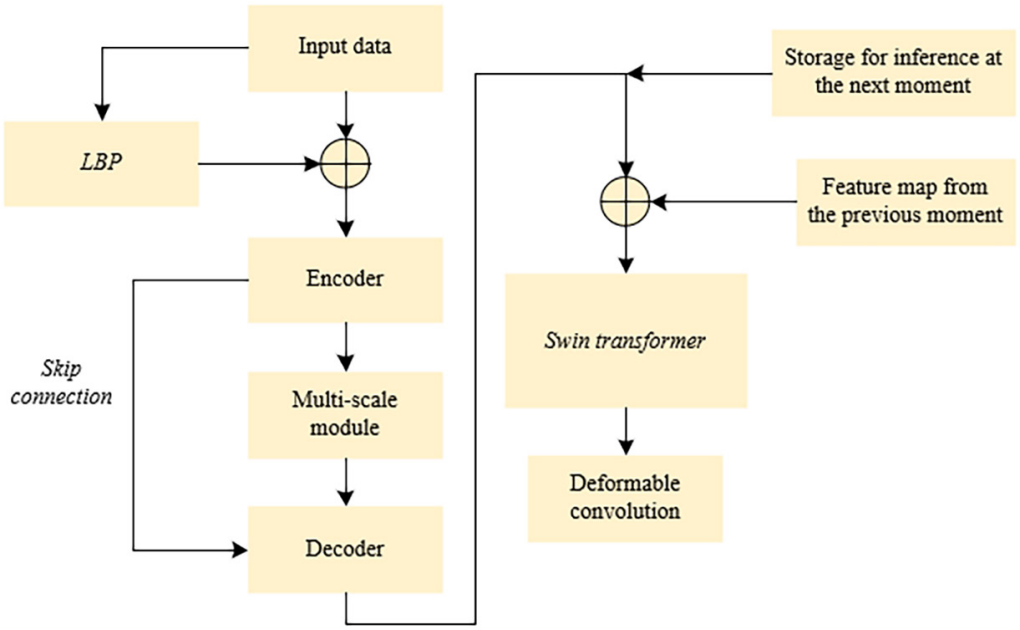


**Fig. 4.** Process of capturing temporal information in satellite remote sensing images

Step 4: Finally, the Swin Transformer is utilized to process and fuse the spatio-temporally aligned satellite remote sensing image feature maps. Specifically, the preprocessed satellite remote sensing images are inputted into the initial layer of the Swin Transformer. At this layer, the model uses sliding windows of various kernel sizes to extract local features. The operation of each window resembles a self-attention mechanism, enabling it to focus on specific parts of the input image and extract relevant features. Subsequent layers of the Swin Transformer can further abstract and refine information from the features extracted by the previous layer. In these layers, the model may use windows of various sizes to capture higher-level features or contextual information at different scales.

After processing through multiple layers, the Swin Transformer integrates features from different levels. This step typically involves the upsampling or downsampling of feature maps and their interconnection. Through this method, the model is capable of integrating multi-scale information, enhancing the representational capability of the features. The core advantage of the Swin Transformer lies in its attention mechanism, which dynamically adjusts the model's focus on different regions within the feature maps. This is particularly important for wildfire risk prediction, as it helps the model identify and emphasize key areas related to fire risk.

Finally, the feature map output by the Swin Transformer can be utilized for subsequent classification, regression, or other tasks. In the context of wildfire risk prediction, the generated feature maps are inputted into the final prediction layer, which forecasts the fire risk level for each area based on the combined features.

The design of the Swin Transformer enables it to process a broad range of contextual information and account for long-distance dependencies during the feature map fusion process. Thus, the model is not only capable of recognizing changes within local areas but also of understanding the correlation of these changes across the entire observation area. This is particularly critical for predicting wildfire risks because various factors within a broader area not only influence local conditions but also affect the occurrence and spread of fires.

## 4 EXPERIMENTAL RESULTS AND ANALYSIS

Table 1 presents a performance comparison of various cloud detection and removal methods across different datasets. Analysis of the data in the table reveals that the algorithm proposed in this study demonstrated significant performance improvements on the MODIS dataset, achieving an accuracy of 48.62%. In comparison, the Fmask algorithm and Sen2Cor achieved accuracies of 22.25% and 22.16%, respectively. Although the recall of the proposed algorithm is relatively lower (46.26%), its global accuracy reached 93.12%, which is significantly higher than the Fmask algorithm's 78.41% and Sen2Cor's 82.36%. On the Landsat dataset, the accuracy (63.24%) of the proposed algorithm was significantly higher than that of the Fmask algorithm (25.36%) and Sen2Cor (27.81%). The proposed algorithm also achieved the highest global accuracy at 95.26%, surpassing the other two methods. When applied to the Sentinel-2 dataset, the proposed algorithm outperformed the convolutional neural networks (CNNs) (41.56%) and SCCN (22.58%) methods in terms of accuracy. In terms of recall, the performance of the proposed algorithm (71.45%) was significantly superior to the other methods, achieving the highest global accuracy as well (92.36%). CNNs showed the highest accuracy (93.21%) and global accuracy (98.56%) on this dataset, but the proposed algorithm excelled in terms of recall at 93.56%. It can be concluded that across all datasets, the algorithm proposed in this study outperformed the comparison methods in terms of accuracy,

especially on the MODIS and Landsat datasets, where the accuracy improvement was particularly notable. The global accuracy, which considers the accuracy across all categories, was highest for the proposed method on all datasets, indicating that this method offered the most reliable cloud detection performance overall. In most cases, an increase in accuracy tends to decrease recall, and vice versa. However, the algorithm proposed in this study significantly improved accuracy while maintaining a relatively high recall, especially on the Sentinel-2 and GOES datasets.

Figure 5 illustrates the variation in the root mean square error (RMSE) of predictions with different feature combinations as the number of training steps increases. It is observed that when only surface temperature is used, the RMSE of predictions starts at 60 and shows a general decreasing trend with an increase in training steps. However, the reduction is limited, eventually stabilizing within the range of 14–17. When surface temperature is combined with the vegetation state, the initial RMSE is similar. However, as training steps increase, the RMSE exhibits a more pronounced downward trend, eventually stabilizing within the range of 13–15. This indicates that including vegetation state significantly improves prediction performance. With the integration of surface temperature, vegetation state, and optical flow motion information, the RMSE decreases significantly from 60 to a range of 9–11. This illustrates that incorporating optical flow motion information enhances prediction accuracy even further. For all feature combinations, RMSE shows a decreasing trend with the increase in training steps, suggesting that prediction capabilities improve as the model training progresses. The most significant decrease in RMSE occurs during the early training steps (0–40 steps), suggesting rapid improvement in predictive ability during the initial learning phase. In subsequent training steps, the rate of decrease slows and eventually stabilizes as the model begins to converge. The algorithm proposed in this paper combines surface temperature, vegetation state, and optical flow motion information. It achieves a significant reduction in RMSE compared to using a single feature or combinations of two features. This reduction shows the efficiency of the algorithm in capturing temporal information. Especially, incorporating optical flow motion information into the prediction model significantly enhances the accuracy of wildfire risk prediction.

**Table 1.** Performance comparison of different cloud detection and removal methods

| Dataset | Neural Network | Accuracy (%) | Recall (%) | Global Accuracy (%) |
|---|---|---|---|---|
| *MODIS* | *Fmask algorithm* | 22.25 | 78.51 | 78.41 |
| | *Sen2Cor* | 22.16 | 81.23 | 82.36 |
| | *Proposed algorithm* | 48.62 | 46.26 | 93.12 |
| *Landsat* | *Fmask algorithm* | 25.36 | 84.25 | 84.12 |
| | *Sen2Cor* | 27.81 | 83.26 | 87.23 |
| | *Proposed algorithm* | 63.24 | 51.24 | 95.26 |
| *Sentinel-2* | *CNNs* | 41.56 | 55.69 | 82.14 |
| | *GANs* | 34.26 | 58.4 | 75.16 |
| | *SCCN* | 22.58 | 35.6 | 71.56 |
| | *Proposed algorithm* | 43.87 | 71.45 | 92.36 |
| *GOES* | *CNNs* | 87.36 | 84.26 | 85.32 |
| | *GANs* | 61.23 | 93.26 | 73.84 |
| | *SCCN* | 91.25 | 77.25 | 92.36 |
| | *CNNs* | 93.21 | 93.56 | 98.56 |

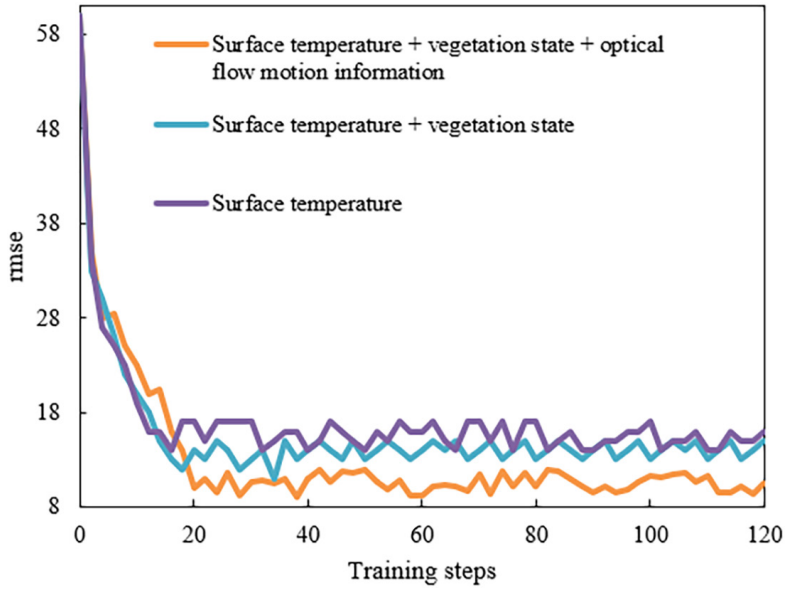


**Fig. 5.** Prediction RMSE curves under different feature map fusion conditions

**Table 2.** Comparison of mean energy values for satellite remote sensing image features

| Mean Energy Value | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| Surface temperature | 0.7541 | 0.8878 | 0.7846 | 0.8745 | 0.6125 |
| Vegetation state | 0.6359 | 0.3365 | 0.3745 | 0.4125 | 0.5748 |
| Fire traces | 0.4215 | 0.7748 | 0.7895 | 0.6592 | 0.5236 |
| Smoke | 0.9236 | 0.8795 | 0.9562 | 0.9421 | 0.9754 |
| Human activities | 0.7895 | 0.8125 | 0.8897 | 0.8895 | 0.6895 |

**Table 3.** Comparison of mean entropy values for satellite remote sensing image features

| Mean Entropy Value | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| Surface temperature | 0.9548 | 0.3526 | 0.7456 | 0.4582 | 1.3625 |
| Vegetation state | 0.9362 | 1.5689 | 1.5623 | 1.3625 | 1.2365 |
| Fire traces | 1.5484 | 0.5896 | 0.6359 | 0.9265 | 1.2458 |
| Smoke | 0.2356 | 0.4123 | 0.0836 | 0.1236 | 0.1326 |
| Human activities | 0.7356 | 0.6895 | 0.4489 | 0.4859 | 0.9354 |

**Table 4.** Comparison of mean inverse difference moment values for satellite remote sensing image features

| Mean Inverse Difference Moment Value | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| Surface temperature | 0.7456 | 0.8875 | 0.7845 | 0.8741 | 0.5894 |
| Vegetation state | 0.6325 | 0.3362 | 0.3745 | 0.4123 | 0.5784 |
| Fire traces | 0.4216 | 0.7784 | 0.7895 | 0.6589 | 0.5231 |
| Smoke | 0.9236 | 0.8845 | 0.9652 | 0.9456 | 0.9451 |
| Human activities | 0.7851 | 0.8256 | 0.8895 | 0.8261 | 0.6895 |

**Table 5.** Performance comparison of different temporal information capture methods for satellite remote sensing images

| Model | Flame Detection | | | Smoke Detection | | | FPS |
|---|---|---|---|---|---|---|---|
| | Precision | Recall | F1-Score | Precision | Recall | F1-Score | |
| *LSTM* | 0.96 | 0.98 | 0.96 | 0.97 | 0.97 | 0.97 | 11.3 |
| *ConvLSTM* | 0.97 | 0.95 | 0.96 | 0.96 | 0.97 | 0.97 | 11.8 |
| *PredRNN* | 0.97 | 0.66 | 0.77 | 0.91 | 0.91 | 0.91 | 11.1 |
| *Autoencoders* | 0.98 | 0.97 | 0.97 | 0.96 | 0.97 | 0.96 | 7.7 |
| *STCN* | 0.91 | 0.95 | 0.91 | 0.87 | 0.93 | 0.92 | 12.3 |
| The proposed method | 0.99 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 11.2 |

The data analysis of the mean energy values, entropy values, and inverse difference moment values, as shown in Tables 2–4, reveals the following insights: Samples 2 and 3 show high energy values and inverse difference moment values for fire traces and surface temperature. Combined with the analysis of entropy values, these samples indicate recent fire activity in the regions. Sample 5 displays a significantly high entropy value for surface temperature. This suggests that fire has affected the area, resulting in an uneven distribution of temperature. High energy values and inverse difference moment values for smoke in sample 5 indicate dense smoke, which is consistent with post-fire conditions. High entropy values for vegetation state, particularly in samples 2 and 3, suggest a significant impact on vegetation or diversity. It can be concluded that the temporal information capturing technique proposed in this paper, through comprehensive analysis of these features, can more accurately identify fire-prone areas, offer early warnings, and support disaster response and management.

Based on the performance comparison of various methods for capturing temporal information in satellite remote sensing images as presented in Table 5, the following analysis is conducted: In flame detection, the method proposed in this paper demonstrates the highest precision (0.99), indicating that almost all areas identified as flames by this method are indeed flames. The long short-term memory (LSTM) model demonstrates the highest recall rate (0.98), indicating its ability to detect the majority of flame areas effectively. The F1-score, which combines precision and recall, is highest for the method proposed in this paper, LSTM, and convolutional LSTM (ConvLSTM) (0.96). This indicates that these models achieve a good balance between accuracy and coverage in flame detection. In smoke detection, the method proposed in this paper, along with LSTM, ConvLSTM, and autoencoders, shows similar performance in precision and recall (0.97). The F1-score for the methods proposed in this paper, LSTM, and ConvLSTM is also 0.97, indicating balanced performance in smoke detection. In terms of frames per second (FPS), the FPS for the method proposed in this paper (11.2) is comparable to LSTM and ConvLSTM, suggesting that it maintains high detection performance without significantly sacrificing processing speed. In summary, the method proposed in this paper offers the highest precision in both flame and smoke detection tasks and matches the best-performing models in terms of F1-score. This indicates that the system can identify fire-related features with high accuracy while maintaining a low false alarm rate. Additionally, its processing speed is comparable to other high-performing models, which is crucial for fire detection because rapid response is key to mitigating fire damage.

**Table 6.** Statistical results of large-scale wildfire risk prediction using multi-color space

| | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| Total frames | 889 | 300 | 1210 | 1578 | 321 |
| Flame frames | 889 | 123 | 298 | 823 | 124 |
| Flame false frames | 0 | 0 | 0 | 0 | 0 |
| Flame miss frames | 3 | 0 | 32 | 75 | 24 |
| Flame false | 0% | 0% | 0% | 0% | 0% |
| Flame accuracy | 98.9% | 100% | 97.1% | 94.5% | 91.2% |
| Flame miss | 0.3% | 0% | 12.1% | 9.2% | 21% |
| Smoke frames | / | / | 93.6 | 1515 | / |
| Smoke false frames | / | / | 25 | 142 | / |
| Smoke miss frames | / | / | 117 | 162 | / |
| Smoke false | / | / | 2.6% | 9.4% | / |
| Smoke accuracy | / | / | 89.1% | 89.5% | / |
| Smoke miss | / | / | 12.3% | 11.2% | / |

Table 6 presents the statistical findings of a novel method for capturing temporal information to predict large-scale wildfire risk across five distinct samples. It is observed that the number of false flame frames is zero across all five samples, indicating no false detections, which is a significant result. Despite a 0% false detection rate, missed detections are still present, particularly in samples 3, 4, and 5, with missed rates of 12.1%, 9.2%, and 21%, respectively. The false alarm rate consistently remains at 0%, demonstrating that in frames without flames, this method does not erroneously detect flames. Flame accuracy is relatively high across all samples, notably reaching 100% for sample 2, while sample 5 has the lowest accuracy at 91.2%. The flame miss rate shows variability, with the highest rate of 21% in sample 5. The number of false smoke frames indicates the presence of false detections, especially in sample 4 at 9.4%. The number of missed smoke frames highlights issues with smoke detection, with a miss rate of 11.2% in sample 4. The false smoke rate is only 2.6% in sample 3. Smoke accuracy in samples 3 and 4 is 89.1% and 89.5%, respectively, indicating relatively high accuracy in smoke detection. The smoke miss rate in both samples exceeds 10%, highlighting the necessity for enhancing smoke detection by the algorithm. Based on the data analysis from Table 6, it can be concluded that the temporal information capture technique proposed in this study performs exceptionally well in flame and smoke detection, offering the highest precision and maintaining high accuracy levels overall. Although some samples exhibit missed detections, the overall accuracy remains at a high level.

## 5 CONCLUSION

This study introduced a cloud detection and removal algorithm based on deep learning aimed at enhancing the clarity and usability of satellite remote sensing images. The algorithm was demonstrated to accurately identify clouds under complex background conditions and effectively eliminate cloud influences from remote sensing images. A novel technique for capturing temporal information was proposed, designed specifically for processing and analyzing vast amounts of remote sensing data. This technique focuses on extracting key features from time series data to construct accurate wildfire risk prediction models.

Experimental results compared the existing methods with the algorithm proposed in this paper, highlighting the advantages of the letter in cloud identification and removal. Plotting the RMSE prediction curve under various feature map fusion conditions confirmed the significance of feature fusion in enhancing prediction accuracy. The study compared the effectiveness of feature extraction by comparing metrics such as mean energy values, mean entropy values, and mean inverse various moment values. The performance of different methods in capturing temporal information demonstrates the effectiveness of the proposed method in extracting time series features. Results of large-scale wildfire risk prediction using multiple color spaces were presented, assessing model performance based on the accuracy of detecting flames and smoke.

Considering the content of the study and the experimental outcomes, it can be concluded that the algorithms and techniques proposed in this paper are innovative and effective for large-scale wildfire risk prediction. They have improved the processing and analysis of remote sensing images, enhancing the accuracy of wildfire risk predictions. Future work may focus on further reducing false and missed detections, as well as expanding applications to predict other types of natural disasters.

## 6 ACKNOWLEDGMENT

The Natural Science Foundation of Heilongjiang Province, China (Grant No.: LH2021D016).

## 7 REFERENCES

[1] L. Collins, R. Trouvé, P. J. Baker, B. Cirulus, C. R. Nitschke, R. H. Nolan, L. Smith, and T. D. Penman, "Fuel reduction burning reduces wildfire severity during extreme fire events in south-eastern Australia," *Journal of Environmental Management*, vol. 343, p. 118171, 2023. https://doi.org/10.1016/j.jenvman.2023.118171

[2] J. Nan, J. Wang, H. Wu, and K. Li, "Optimized extreme learning machine by an improved Harris Hawks optimization algorithm for mine fire flame recognition," *Mining, Metallurgy & Exploration*, vol. 40, no. 1, pp. 367–388, 2023. https://doi.org/10.1007/s42461-022-00719-5

[3] D. Avetisyan, E. Velizarova, and L. Filchev, "Post-fire forest vegetation state monitoring through satellite remote sensing and in situ data," *Remote Sensing*, vol. 14, no. 24, p. 6266, 2022. https://doi.org/10.3390/rs14246266

[4] A. Sukmono, F. Hadi, E. Widayanti, A. L. Nugraha, and N. Bashit, "Identifying burnt areas in forests and land fire using multitemporal Normalized Burn Ratio (NBR) index on Sentinel-2 Satellite imagery," *International Journal of Safety and Security Engineering*, vol. 13, no. 3, pp. 469–477, 2023. https://doi.org/10.18280/ijsse.130309

[5] V. Semeniaka, V. Zatserkovny, A. Ilchenko, P. Trofymenko, and O. Nikolaienko, "Application of GIS, remote sensing and GPS technologies for forest fire monitoring tasks," in *15th International Conference Monitoring of Geological Processes and Ecological Condition of the Environment*, 2021, vol. 2021, no. 1, pp. 1–5. https://doi.org/10.3997/2214-4609.20215K2062

[6] A. Sagynbayeva, B. Mambetov, A. Kalachev, Z. Boranbay, Y. Borissenko, Z. Baigazakova, and K. Turlybekov, "Assessing pine forest restoration post-fire using NDVI and GIS technologies," *International Journal of Design & Nature and Ecodynamics*, vol. 18, no. 6, pp. 1459–1467, 2023. https://doi.org/10.18280/ijdne.180620

[7] J. Deng, J. L. T. Cullen, Y. Xue, F. Zhou, and B. Shi, "Spatial-temporal analysis of coal fire risk identification and suppression assessment with satellite time series mapping 2013–2020 in Midong coalfield, Xinjiang, China," *International Journal of Remote Sensing*, vol. 44, no. 7, pp. 2236–2272, 2023. https://doi.org/10.1080/01431161.2023.2197133

[8] X. Chen, T. Li, L. Ruan, K. Xu, J. Huang, and Y. Xiong, "Research and application of fire risk assessment based on satellite remote sensing for transmission line," in *Proceedings of the World Congress on Engineering and Computer Science*, San Francisco, CA, USA, 2015, pp. 284–287.

[9] H. I. M. Al-Hilfi and M. M. J. Al-Nayar, "Increase the WSN-lifespan used in monitoring forest fires by PSO," *Mathematical Modelling of Engineering Problems*, vol. 9, no. 3, pp. 583–590, 2022. https://doi.org/10.18280/mmep.090304

[10] W. Chen, Y. Zhou, E. Zhou, Z. Xiang, W. Zhou, and J. Lu, "Wildfire risk assessment of transmission-line corridors based on naïve Bayes network and remote sensing data," *Sensors (Switzerland)*, vol. 21, no. 2, pp. 1–16, 2021. https://doi.org/10.3390/s21020634

[11] R. Lasaponara, A. Aromando, G. Cardettini, and M. Proto, "Fire risk estimation at different scales of observations: An overview of satellite based methods," in *Computational Science and Its Applications–ICCSA 2018: 18th International Conference*, Melbourne, VIC, Australia, 2018, pp. 375–388. https://doi.org/10.1007/978-3-319-95174-4_30

[12] J. Liu, W. Hou, X. Luo, J. Su, Y. Hou, and Z. Wang, "SI-SA GAN: A generative adversarial network combined with spatial information and self-attention for removing thin cloud in optical remote sensing images," *IEEE Access*, vol. 10, pp. 114318–114330, 2022. https://doi.org/10.1109/ACCESS.2022.3213354

[13] X. Xie, X. Hou, and L. Cao, "Supervised cloud presence detection in remote sensing image," in *3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, 2021, pp. 124–127. https://doi.org/10.1109/IAECST54258.2021.9695686

[14] H. S. Na and C. B. Chen, "Deep learning methods for road extraction from remote sensing imagery," in *IIE Annual Conference. Proceedings*, 2022, pp. 1–6.

[15] H. Wang, F. Bu, J. Wen, J. Yang, W. Yao, and Q. Liu, "Time series dynamic changing analysis on the mine geological environment in a natural reserve in Tibet based on the satellite remote sensing technology," in *International Conference on Geographic Information and Remote Sensing Technology (GIRST 2022)*, 2023, vol. 12552, pp. 208–214. https://doi.org/10.1117/12.2667528

[16] C. Zhong, F. Bu, N. Wang, J. Wen, W. Lu, and L. Wu, "Study on estimation of water changes in plateau lake based on long time series remote sensing data: Takes Selinco Lake as an example," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2023, vol. 12617, p. 126172L. https://doi.org/10.1117/12.2664929

[17] J. Zhou, M. Menenti, L. Jia, B. Gao, F. Zhao, Y. Cui, X. Xiong, X. Liu, and D. Li, "A scalable software package for time series reconstruction of remote sensing datasets on the Google Earth Engine platform," *International Journal of Digital Earth*, vol. 16, no. 1, pp. 988–1007, 2023. https://doi.org/10.1080/17538947.2023.2192004

## 8 AUTHORS

**Boxin Li** is affiliated with the College of Computer and Control Engineering at Northeast Forestry University, located in Harbin, China (E-mail: liboxin15645101512@163.com).

**Hong'e Ren** is a member of the College of Computer and Control Engineering at Northeast Forestry University in Harbin, China (E-mail: rhe@nefu.edu.cn).

**Jing Tian** is with the College of Surveying and Mapping Engineering at the Heilongjiang Institute of Technology, Harbin, China (E-mail: tianjing@hljit.edu.cn).