Deep Learning-based Detection of Feature Changes in Arable Land

Ming Gao [1], Can Ma [1], Qi Liu [1], Zhaoyuan Yu [1,2,3], Linwang Yuan[1,2,3], Wen Luo [1,2,3,*]

[1] School of Geography, Nanjing Normal University, Nanjing, 210023, China.
[2] Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China.
[2] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, 210023, China.

*Correspondence: luowen@njnu.edu.cn

Received: 15 October 2022; Revised: 8 March 2023; Accepted: 16 March 2023; Published 31 March 2023

Abstract: In economically developed regions, the occupation of arable land has long been a significant concern, and identifying cultivated land accurately is crucial for mitigating this issue. To address the problem of "pseudo-change" caused by complex backgrounds, uneven brightness, and natural growth, we propose a change detection algorithm based on twin neural networks. Our algorithm comprises two sub-networks and a decision layer that use a multi-scale feature model and an improved spatio-temporal attention module to extract and process features from two temporal images. We validate our algorithm using high-resolution UAV aerial photography data with a resolution of 0.3m and 0.05m. Our proposed method achieves an F1 coefficient of over 87% in the cultivated land feature change detection dataset, which is more than 45% higher than the traditional STANet. The contribution of our research lies in developing a more effective method for identifying cultivated land in economically developed regions, which can help mitigate the serious "man-made disaster" of arable land occupation. Our proposed method has the potential to contribute significantly to the field of precision agriculture and land resource management.

Keywords: Deep Learning; Arable Land Protection; Remote Sensing Identification; Pseudo-change

1.0 Introduction

The planet is facing severe disasters due to factors such as climate change and environmental damage. Arable land, one of the most important natural resources for human beings, is also threatened by disasters. Arable land protection is an important basis for food security, and governments and researchers have conducted a series of studies on arable land protection, which has always been a hot topic of research in agronomy, ecology and geography (Hu et al., 2021).

In contrast to natural disasters such as floods, extreme weather, pests and diseases, arable land in more economically developed regions also faces the "man-made disaster" of arable land encroachment. For a long time, the authorities at all levels have been strictly prohibiting and cracking down on the illegal occupation and destruction of arable land and other issues related to agriculture, but due to the huge profit motive and the complexity of the situation involved, such cases have been prohibited but not stopped. For the encroachment of arable land, the main research currently focuses on the detection of buildings in arable land (He et al., 2016). Traditional building detection results mainly come from external mapping and visual interpretation, which are affected by topography, weather and other environmental factors, and are time-consuming and difficult to ensure their timeliness; visual interpretation methods to obtain building information on high-precision remote sensing images are time-consuming and costly. Since entering the automated building detection stage, researchers have used computer technology to achieve the extraction of building targets from light detection and ranging （LiDAR) data and remotely sensed images by identifying building features such as building edge contours and corner point information (Jin et al., 2013). Using supervised classification methods combining classifiers such as back propagation neural networks, support vector machines and random forests (Liu et al., 2015) to produce sample datasets and train the classification models, the classification accuracy was effectively improved. However, due to the diverse building morphology, different angles of remote sensing image photography and resolution, the trained models generally have the deficiency of poor generalization ability. With the introduction of convolutional neural networks (CNN), this situation has been largely improved and the models have a strong generalisation capability. Based on this idea of end-to-end convolutional neural networks, U-Net and mask regional convolutional neural network (Mask RCNN) (Chen et al., 2018) were derived to further improve the performance of building extraction models. The proposed twin neural networks have further promoted the development of remote sensing image change detection. At present, the detection of arable land features is only limited to buildings, but in fact, the risk of "non-food" exists in structures such as sheds, which should also be the target of arable land conservation feature detection, but there is less comprehensive research in this area (Gao et al., 2022).

In addition, due to the special characteristics of cropland, cropping patterns and the natural growth of plants affect the external characteristics of cropland, different crops are grown at different times of the year, and the colour, volume and form of plants in cropland vary from season to season, making it difficult to detect changes. High-resolution remote sensing data provides a rich data source for change detection, and the features in high-resolution data are richer and more detailed, but this also brings more "pseudo-change" information to the change detection work, which affects the accuracy of change detection (Beke et al., 2021).

The aim of this study is to address the problems of complex background, uneven brightness and "pseudo-change" caused by natural growth factors in arable land areas. The algorithm makes full use of the twin neural network to detect the change in the area. The algorithm gives full play to the feature mining ability of twin neural networks, which can effectively extract the change detection information of cropland features and improve the detection accuracy of cropland change detection.

2.0 Study Area

Jintan District, which belongs to Changzhou City, Jiangsu Province, is located in the south of Jiangsu Province, between 31°33′42″-31°53′22″N and 119°17′45″-119°44′59″E. In this paper, a designated area in Jintan District was selected as the experimental area, which has a strong agricultural base, a large proportion of agricultural land, intensive human activities in the surrounding area, and the presence of arable land occupation, and is suitable for arable land feature change detection study.

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

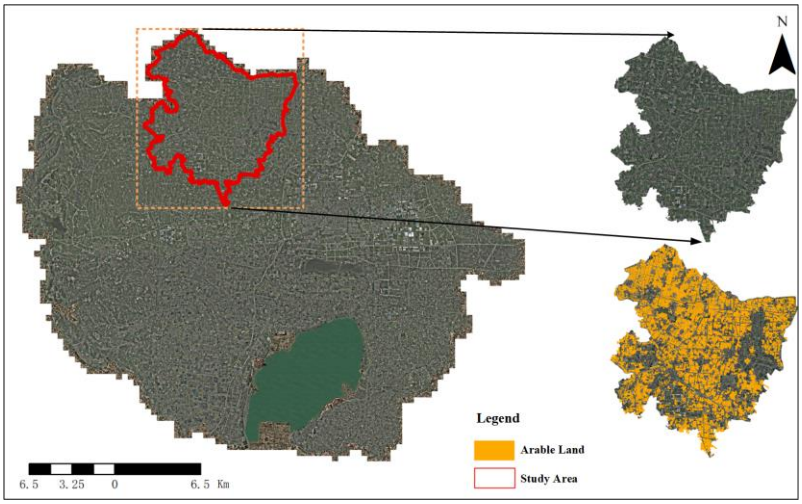


Figure 1: Study area location map.

3.0 Materials and Methods

3.1 Data

The data in this study are mainly remote sensing images in two-time phases: UAV remote sensing data for 2018 and 2021, Remote sensing data for 2018 is 0.05m in summer, remote sensing data for 2021 is 0.3m in spring.

3.1 3.2 Methodology

The proposed twin neural network for arable land feature change detection focuses on the Spatial-Temporal Attention Neural Network (STANet) with improvements in the network structure of the feature extraction layer and for the spatio-temporal attention module.

3.2.1 *STANet basic network architecture*

STANet is designed to model the spatio-temporal relationships with a self-attentive module that calculates the attention weight between the same positions of two temporal images and uses this weight to generate saliency features. In addition, the model partitions the image into multi-scale sub-regions so that spatio-temporal dependencies can be established at different scales and can be adapted to various sizes of images (H. Chen & Shi, 2020). The basic structure of the STANet model is shown in Figure 2.

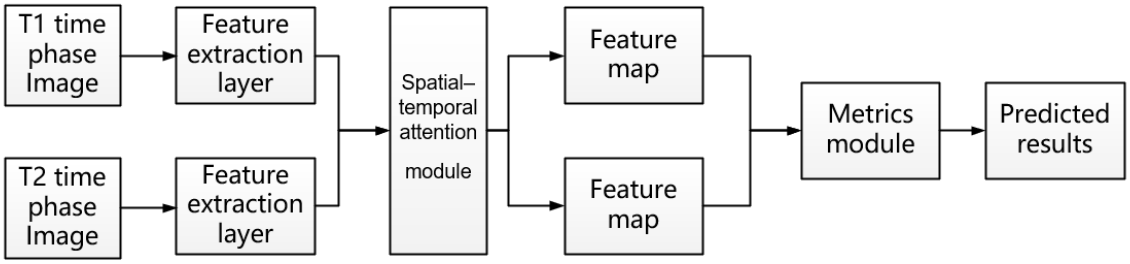


Figure 2: Schematic diagram of the STANet structure.

3.2.2 *Improvement of the network structure of the feature extraction layer*

Feature extraction is the most important part of change detection work. In traditional convolutional neural networks, the input results of this layer are derived from the output results of the previous layer. As the number of layers increases, the network can perform more complex feature extraction, but the network suffers from degradation, i.e. the accuracy of the network decreases when the depth of the network increases. Unlike traditional convolutional neural networks where the parameter layer learns the mapping between input and output, the parameter layer in ResNet networks learns the residual representation of the input and output, i.e. the residual problem is used to solve the model training degradation problem (Wu et al., 2019).

Let the features that can be learned for a stacked-layer structure when the input is x be $H(x)$, whose residuals can be expressed as Eq.1.

$$F(x) = H(x) - x \quad (1)$$

Then the original learned features can be expressed as $F(x) + x$. Figure 3 shows the residual basic unit. According to the figure, it can be seen that when the input features x satisfy the criteria, $H(x)$ and x are similar, then it means that F(x) tends to 0 and x will be forward propagated in a constant mapping and its output will not affect the next layer.

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

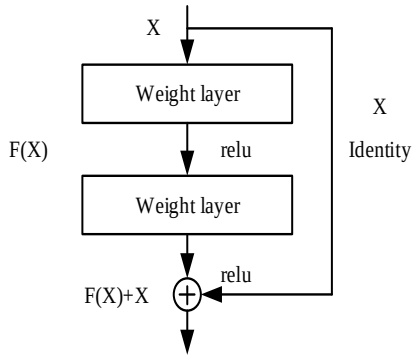


Figure 3: Basic residual unit

In the field of computer vision, multi-scale features are crucial. The multi-scale feature model (Res2Net) (Chen et al., 2018) is a successful variant of ResNet that successfully introduces the scale problem into ResNet, allowing it to be integrated with other modules without increasing the amount of computational load. Res2Net uses a variable structure to replace the convolutional kernel in the original ResNet, which implements layered connections in a separate residual block that embodies multi-scale features. In addition, Res2Net increases the perceptual domain of each network layer and does not increase the network parameters. Such features allow Res2Net to be inserted into other convolutional neural networks without increasing the computational load while improving network performance. The building blocks of Res2Net are shown in Figure 4.

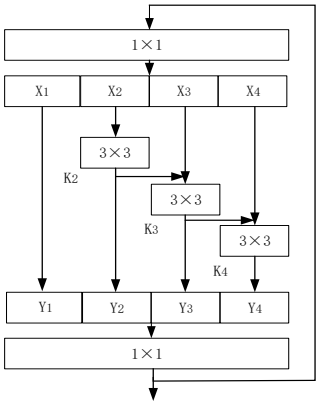


Figure 4: Schematic diagram of the Res2Net structure block

In previous research by scholars, the detection results of Res2Net on several datasets such as ImageNet and CIFAR-100 are better than those of ResNet. In this paper, in order to further improve the extraction accuracy of the feature extraction layer, Res2Net is used as the extraction of the feature extraction layer to effectively improve the extraction accuracy of the feature extraction layer, and thus improve the change detection detection effect (Zhang et al., 2020).

3.2.3*Improvement of basic spatial–temporal attention module*

The Basic Spatial-temporal Attention Module (BAM) is a combination of the self-attentive mechanism in natural language processing and convolutional neural networks (W. Du et al., 2018). The computational flow of BAM is shown in Figure 5.

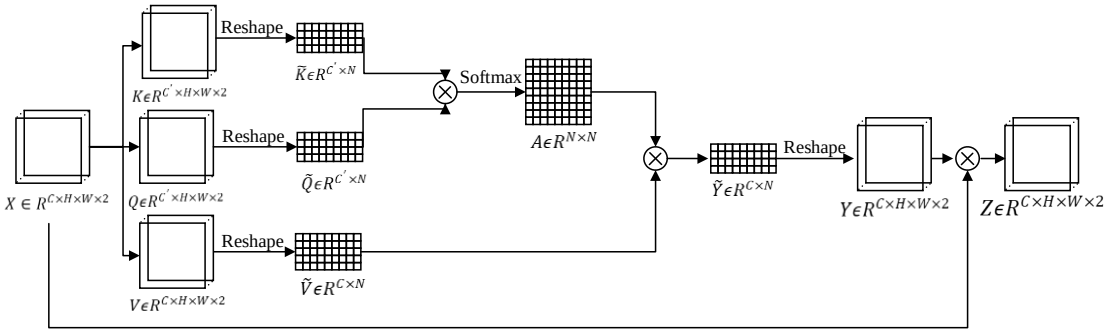


Figure 5: Schematic diagram of BAM

Let the length of the image be H, the width be W, the number of convolution layers be C, the feature tensor be X, $X \in R^{C\times H\times W\times 2}$, and X can be obtained as Key, $Query$ and $Value$ by three different convolution layers, i.e., three feature maps are obtained as K, Q and V, where $K\epsilon R^{C'\times H\times W\times 2}$, $Q\epsilon R^{C'\times H\times W\times 2}$, $V\epsilon R^{C\times H\times W\times 2}$. To reduce the number of network parameters, K, Q and V are resampled to obtain $\tilde{K}$, $\tilde{Q}$, $\tilde{V}$, and such that X and the output feature tensor is Z. The computational steps are shown below:

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

Step 1: Computing the transpose of $\tilde{K}$ and the dot product of $\tilde{Q}$ and dividing by $\sqrt{C'}$.

Step 2: Normalisation by the softmax function to obtain the weight coefficient A, calculated as in equation 2.

$$A = softmax(\frac{\tilde{K}^T\tilde{Q}}{\sqrt{C'}}) \quad (2)$$

Step 3: The weighting coefficients A and Value are weighted and summed to obtain, as in equation 3

$$\tilde{Y} = \tilde{V} * A \quad (3)$$

Step 4: Resampling $\tilde{Y}$ to obtain Y, $Y \epsilon R^{C\times H\times W\times 2}$.

Step 5: Calculating the output feature tensor Z, $Z \epsilon R^{C\times H\times W\times 2}$ by equation 4.

$$Z = Y + X \quad (4)$$

The above steps result in a feature tensor Z output by the self-attentive module, which takes into account the influence of any other position on this map. As a result, the BAM module allows the light invariance between pixels to be obtained, effectively improving the alignment robustness of the model.

In traditional pooling operations, the main purpose is to extract invariant features in the image, without focusing on the information that is missing for spatial locations. The Pyramid spatial-temporal attention module (PAM) is a spatial-temporal self-attentive module that aggregates multi-scale information to address this problem. The PAM achieves a more accurate output feature map by extracting the context.

In PAM, it mainly utilises the idea of the pyramid pooling module in the Scene Level Global Information Extraction Model ——Pyramid Scene Parsing Network(PSPNet), and then implements the design of PAM by combining it with the self-attention mechanism. the computational flow of PAM is shown in Figure 6.

Four different branches are designed in the PAM, each with a fixed sub-region size set to $s \times s$, $s = \{1,2,4,8\}$, and similar to the BAM, the length of the image is set to H, the width to W, the feature tensor to X, $X \in R^{C\times H\times W\times 2}$, and the output feature tensor to Z. The computational steps are shown below:

Step 1:divide X into $s \times s$ subregions, where each subregion $X_{ij} \in R^{C\times\frac{H}{2}\times\frac{W}{2}\times 2}, i,j = \{1,2\}$.

Step 2: Repeat steps 1-5 in the BAM module for each subregion in each of the four branches to obtain the local attention features at that scale.

Step 3: stitch the local attention on the four branches and then pass the 1×1 convolution layer to obtain the updated residual feature vector Y, $Y \in R^{C\times H\times W\times 2}$.

Step 4: Based on the residual network properties, the output feature tensor Z, Z is obtained according to Equation 5. $Z \in R^{C\times H\times W\times 2}$

$$Z = Y + X \quad (5)$$

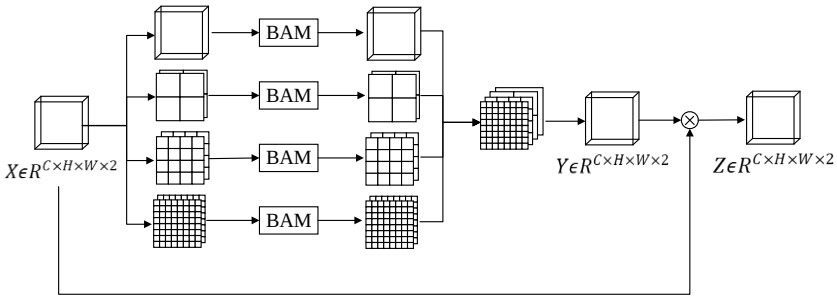


Figure 6: Schematic diagram of PAM

Through the calculation, Z is the final multi-scale attentional feature. The PAM module enables better global contextual information to be obtained, and the four branches fuse the results of local attention operations on different sub-regions, ensuring that each pixel in the image undergoes self-attention operations at different scales, which allows edge information to be preserved for better extraction results. Figure 7 shows the Improved Spatial-temporal Self-Attention Module.

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

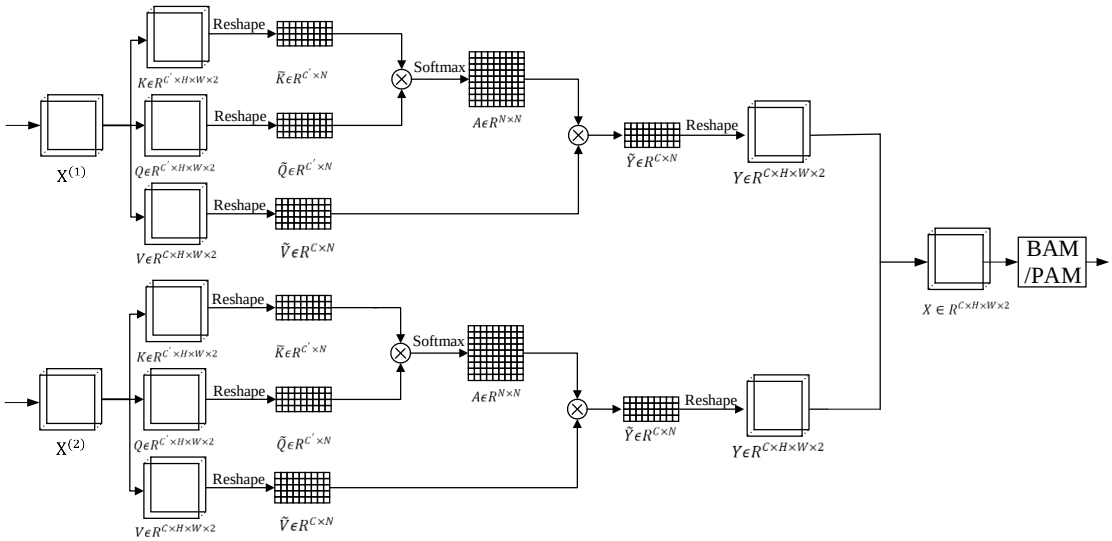


Figure 7: Schematic diagram of the Improved Spatial-temporal Self-Attention Module

3.2.4 *Algorithm for detecting changes in arable land features based on multi-scale features and an improved spatio-temporal self-attentive module*

In order to further improve the recognition capability of the feature extraction layer and the ability to capture distinguishing features, and further reduce the impact of "pseudo-change", this paper proposes a change detection algorithm (STANet_R2N_SAS) based on a multi-scale feature model and an improved spatio-temporal self-attention mechanism module. It first uses the scale features of Res2Net to obtain different scale detail features; then the feature map is processed by the improved spatial attention module of the spatio-temporal attention module, and the double attention mechanism is used to give more weight to the change region, which results in significant changes, and the loss function is used to equalize the change information. The specific network structure is shown in Figure 8.

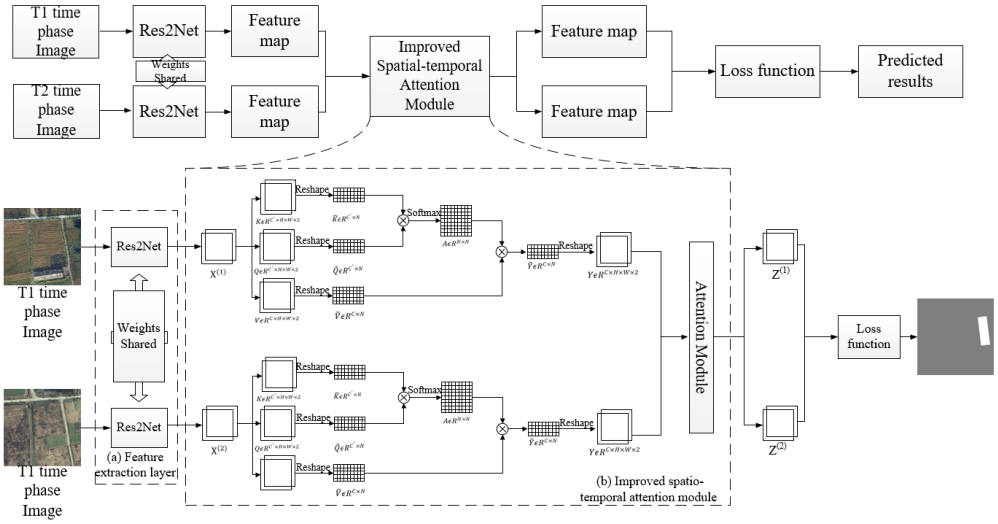


Figure 8: Structure of change detection algorithm based on multi-scale

4.0 Results

4.1 Feature importance

Based on the preliminary analysis of the development law of landslide in the study area, the considered parameters related to landslides in the region are elevation, slope, aspect, distance from road (DFRO), distance from river (DFRI), distance from structure (DFS), vegetation cover and engineering rock group (ERG). The information value model is adopted, and the contribution of various factors to landslide is considered. The information value model is established by using Formula (1) to obtain the predicted information value of each factor, as shown in Table 1, and get the predicted information value of all factors, as shown in Fig.2a. According to the statistics of the distribution data of grid information in the whole area, the abrupt change point of the data is taken as the critical value of grade division, and the vulnerability of landslides in the whole area is divided into four grades: very high susceptibility area, high susceptibility area, medium susceptibility area, and low susceptibility area, as shown in Fig.2b.

In this experiment, the two temporal images were subjected to change detection according to the change detection process of arable land features from high resolution images, using the STANet_R2N_SA-PAM model, and the total area of the final change detection results was

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/). E-ISSN: 2948-5231

0.36 Square kilometres. Figure 9 shows the distribution of arable land feature change information in this region after the arable land change detection process.

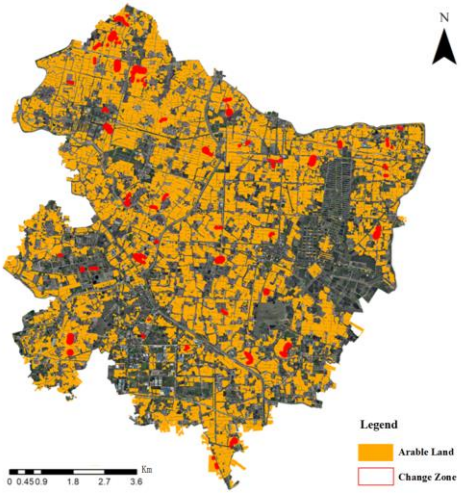


Figure 9: Results of the distribution of change detection results in the experimental area

5.0 Discussion

5.1 Change detection results based on traditional methods

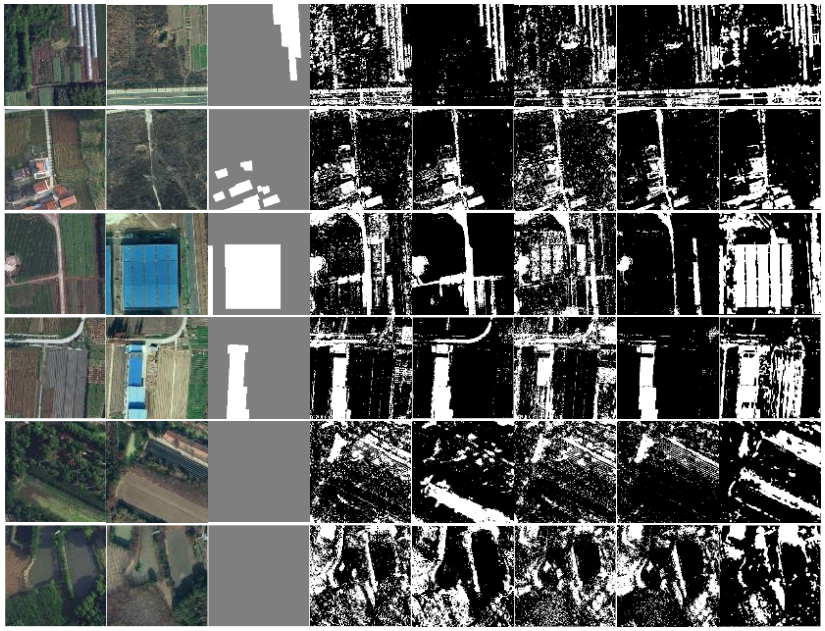
Based on the experimental validation conducted in this subsection, we present Table 1, which showcases various change detection methods, including their traditional counterparts and their deformations. The table classifies the change detection algorithms into three primary categories based on their algorithmic cores, encompassing a total of five detection methods, namely the Slow Feature Analysis Algorithm (SFA) (Shang et al., 2018), Affine Slow Feature Analysis Algorithm (ISFA) (Das et al., 2019), Multivariate Change Detection Algorithm for Typical Correlation Analysis (MAD) (B. Du et al., 2018), Multivariate Change Detection Algorithm for Affine Typical Correlation Analysis (IRMAD) (Jabari et al., 2019), and Principal Component Analysis Clustering Change Detection Algorithm (PCAKmeans) (Leichtle et al., 2017).

Table 1: Change detection methods based on traditional methods and their deformations

Algorithm core	Name
Slow Feature Analysis (SFA)	Change detection method based on slow feature analysis (SFA)
	Change detection method based on iterative slow feature analysis (ISFA)
Typical correlation analysis (MAD)	Multivariate change detection method based on typical correlation analysis (MAD)
	Iterative multivariate change detection based on typical correlation analysis (IRMAD)
K-means clustering, Principal component analysis	Change detection method based on K-means clustering principal component analysis (PCAK-means)

The arable land feature dataset's change detection results using the traditional method are presented in Figure 10. The figures depict the T1 time-phased image, T2 time-phased image, and true change information in 10(a), 10(b), and 10(c), respectively. Upon examining the detection results in Figures 4-5, it becomes apparent that conventional change detection algorithms based on image elements face a common challenge, rendering the detection results neither good nor bad. While change information is extracted in the change detection results, the features' change information cannot be accurately highlighted. From the results' perspective, traditional change detection algorithms are highly susceptible to vegetation change, and the outcomes often display 'pretzel noise,' leading to inconspicuous change information extraction that is mostly irrelevant to the detection's content. Among the five methods selected in this section, the ISFA (Figure 10(e)) and IRMAD (Figure 10(g)) algorithms exhibit clearer boundaries than the other algorithms, with fewer detection result noises and the ability to extract relatively complete change information.

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 10: Comparison results based on conventional methods (a) T1 temporal image; (b) T2 temporal image; (c) label; (d) SFA; (e) ISFA; (f) MAD; (g) IRMAD; (h) PCAKmeans

5.2 Change detection results based on twin neural networks

In order to further evaluate the superiority of the network design performance of this paper, the adaptability and effectiveness of the improved algorithms and improved modules for the detection of change in cultivated land features, this section will design different algorithm modules to evaluate the accuracy of change detection in cultivated land features. The twin network-based change detection algorithms and their deformations selected in this section are shown in Tables 4-4, which classify the change detection algorithms into four categories according to the feature extraction layer network, and refine the algorithms into ten detection methods according to the different attention modules, i.e. by combining different feature extraction layer networks with different attention modules to design ablation experiments, the improved modules proposed in this paper and the improved algorithms on the cropland feature change detection dataset change detection dataset for validation.

In Table 2, where ResNet and Res2Net are different feature extraction layer networks, BAM and PAM are different spatio-temporal attention modules, and SA_BAM and SA_PAM are different improved spatio-temporal attention modules. Programme 3 in Table 4-3 is the twin neural network-based algorithm for detecting changes in cultivated land features proposed in this paper, and the rest are experimental comparison schemes.

Table 2: The hazard information value of factors Change detection methods based on twin networks and their deformations

Design programme	Algorithm name	Feature extraction layer		Attention module		Improved attention module	
		ResNet	Res2Net	BAM	PAM	SA_BAM	SA_PAM
STANet	STANet	√					
		√		√			
		√			√		
Programme1	STANet_SA	√				√	
		√					√
Programme2	STANet_R2N		√				
			√	√			
			√		√		
Programme3	STANet_R2N_SAS		√			√	
			√				√

Table 3 shows the accuracy evaluation results of the deep learning-based change detection model on the cropland feature change detection dataset. The table lists the overall accuracy, precision, recall and F1 coefficient related evaluation metrics for all the deep learning methods. The comparison of the change detection accuracies of the different methods in Tables 4-5 shows that Scenario 1, Scenario 2 and Scenario 3 have significantly improved in each evaluation metric compared to STANet. Scheme I show an overall improvement of more than 0.1 in the F1 coefficient compared to STANet, which indicates the positive effect of the improved spatio-temporal self-attentive module in processing change information. Scenario 2 shows an overall improvement of around 0.3 compared to STANet, which indicates the effective improvement in the model's ability to extract change information after replacing the feature extraction layer. Scenario three shows an overall improvement of more than 0.4 in the F1 coefficient compared to STANet, which illustrates the effectiveness of the proposed algorithm for detecting change in cultivated land features. In addition, Scheme III in turn outperformed the other schemes in other indicators. Therefore, it is concluded from this data set that the design of Scheme III outperforms both Scheme II and Scheme I. Among all the model results, the STANet_R2N_SA-PAM model is the best solution among all the solutions, with the highest F1 coefficient of 0.86 among all the algorithms, and

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

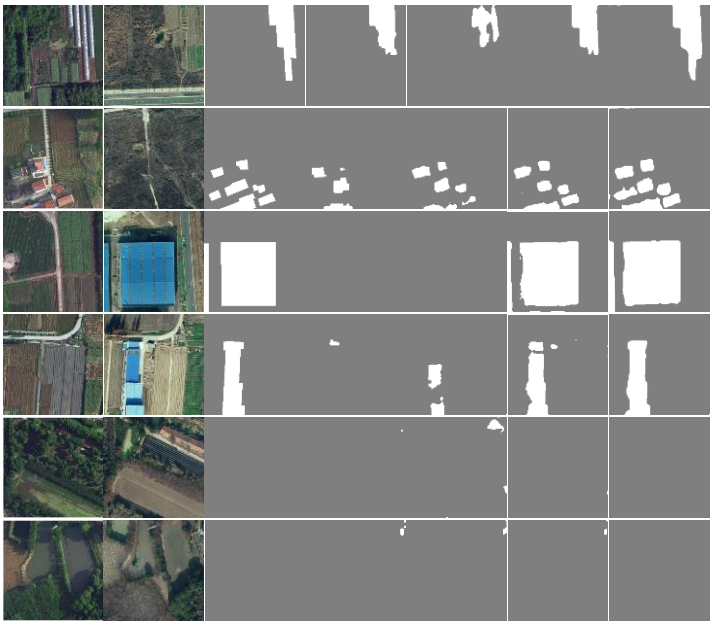
the algorithm obtained the closest results to the ground reference data, thus illustrating the effectiveness of the improved strategy and algorithm proposed in this paper from the data level.

Table 3: The hazard information value of factors Change detection methods based on twin networks and their deformations

Design programme	Feature extraction layer	Attention module	Overall accuracy	Accuracy	Recall rate	F1 coefficient
STANet	ResNet	-	0.976	0.242	0.322	0.283
		BAM	0.976	0.293	0.482	0.362
		PAM	0.989	0.747	0.312	0.438
Programme1 STANet_SA	ResNet	SA_BAM	0.987	0.718	0.317	0.439
		SA_PAM	0.986	0.552	0.569	0.561
Programme2 STANet_R2N	Res2Net	-	0.990	0.616	0.914	0.739
		BAM	0.992	0.718	0.792	0.754
		PAM	0.989	0.682	0.797	0.734
Programme3 STANet_R2N_SAS	Res2Net	SA_BAM	0.991	0.743	0.759	0.756
		SA_PAM	0.996*	0.794*	0.949*	0.871*

In order to visually verify the effect of different models on the extraction of change areas, some typical features are selected for further analysis and evaluation in this paper, and the models are presented in categories according to the attention module.

Figures 11 show the change detection results based on the PAM module for the arable dataset. From the experimental results, the pixel-level change detection method can basically detect most of the change information of the cultivated land features, but the results of the STANet-PAM algorithm in Figure11(d) have more miss detection than other methods, failing to accurately identify the change contours, and there are also cases of false detection, marking unchanged areas in the image as changed areas. The STANet_SA-PAM algorithm in Figure 13(e) is a further improvement of the STANet algorithm. Using the improved PAM module, the method improves the accuracy of detection at the data level compared to the STANet algorithm, but still suffers from incomplete change detection, unclear contours and false detection. The STANet_R2N-PAM algorithm in Figure 13(f) replaces the ResNet in STANet and uses Res2Net as the feature extraction layer network. Compared with the previous two methods, this method can effectively extract most of the change information and can better maintain the feature contour information, but there are still problems of missed detection and false detection. Figure 11(g) the STANet_R2N_SA-PAM algorithm integrates the improved strategies of the first two methods and fuses all the extracted spatial information as far as possible, the results obtained by it are closest to the ground reference data, the accuracy of the feature extraction layer is improved to effectively improve the detection accuracy in the subsequent processing, the refined processing of the improved spatio-temporal attention module suppresses the noise problem The improved refinement of the spatio-temporal attention module suppresses the noise problem and effectively improves the change detection accuracy. The algorithm maintains the feature contour information more completely and effectively eliminates most of the false and missed detections.



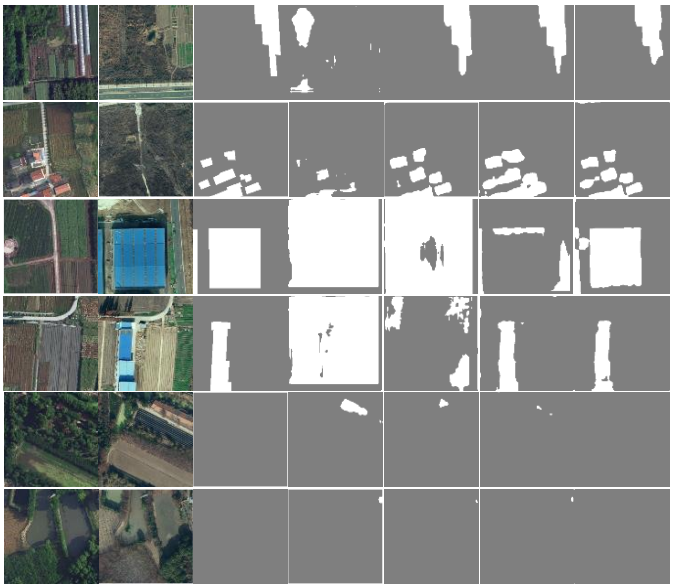
(a) (b) (c) (d) (e) (f) (g)

Figure 11: Comparison results based on PAM modules (a) T1 time-phase image; (b) T2 time-phase image; (c) label; (d) STANet-PAM; (e) STANet_SA-PAM; (f) STANet_R2N-PAM; (g) STANet_R2N_SA –PAM

Figure 12 depicts the change detection outcomes based on the BAM module for the cropland feature dataset. Although the change detection results based on the BAM module are inferior to those based on the PAM module, they can still confirm the efficacy of the algorithmic approach proposed in this paper. Notably, the STANet-BAM algorithm in Figure 12(d) exhibits more false detections than other methods, while the detected change information presents a wide range, incomplete contours, and a large number of unchanged areas labeled as change areas.

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231

Similarly, the STANet_SA-BAM algorithm in Figure 12(e) exhibits comparable issues to the STANet algorithm when using the improved BAM module. However, this approach can effectively detect some change information in certain prediction cases, enhancing detection accuracy and minimizing missed detections. In contrast, the STANet_R2N-BAM algorithm in Figure 12(f), which employs Res2Net as the feature extraction layer network, can extract change information effectively while preserving the feature contour information better than the preceding two methods. Nevertheless, it still suffers from numerous missed detection instances. Finally, the STANet_R2N_SA-BAM algorithm in Figure 12(g) integrates all the spatial information extracted as much as possible, thereby effectively extracting more change information than the preceding methods while maintaining more complete feature contour information. Nonetheless, this algorithm still experiences some false detections, and the extraction of target change information is not as effective as that of the PAM module algorithm.



(a) (b) (c) (d) (e) (f) (g)

Figure 12: Comparison results based on BAM modules (a) T1 time-phase image; (b) T2 time-phase image; (c) label; (d) STANet-BAM; (e) STANet_SA-BAM; (f) STANet_R2N-BAM; (g) STANet_R2N_SA -BAM

6.0 Conclusions

In this paper, a twin neural network-based change detection method is designed to detect the change of cultivated land features in two temporal images. The results show that the algorithm fully exploits the feature mining capability of twin neural networks and can effectively extract the change detection information of cultivated land features. In order to further verify the recognition and generalisation capabilities of the algorithm, the algorithm was evaluated at two levels: the source of the algorithm and the source of the dataset. The results show that the deep learning-based change detection algorithm can, to a certain extent, overcome the influence of "pseudo-change" information and extract the change detection information of interest, with better results than the change detection algorithm based on traditional methods.

In the future, the detection algorithm can be further optimised to add more types of arable land features to support the enforcement and supervision of the supervisory authorities on the indiscriminate occupation of arable land, so as to timely detect the illegal acts, mitigate the "man-made disasters" faced by arable land and effectively protect food security.

Acknowledgement: This work is supported in part by the National Natural Science Foundation of China (No. 42230406, 41976186)

Conflicts of Interest: The authors declare no conflict of interest.

References

Beke, Á. K., Gyürkés, M., Nagy, Z. K., Marosi, G., & Farkas, A. (2021). Digital twin of low dosage continuous powder blending – Artificial neural networks and residence time distribution models. *European Journal of Pharmaceutics and Biopharmaceutics, 169*, 64–77. https://doi.org/10.1016/j.ejpb.2021.09.006

Chen, G., Zhang, X., Wang, Q., Dai, F., Gong, Y., & Zhu, K. (2018). Symmetrical Dense-Shortcut Deep Fully Convolutional Networks for Semantic Segmentation of Very-High-Resolution Remote Sensing Images. IEEE Journal of Selected Topics in *Applied Earth Observations and Remote Sensing, 11*(5), 1633–1644. https://doi.org/10.1109/JSTARS.2018.2810320

Chen, H., & Shi, Z. (2020). A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sensing, 12*(10), 1662. https://doi.org/10.3390/rs12101662

Das, S., Chakraborty, S., Routray, A., & Deb, A. K. (2019). Fast Linear Unmixing of Hyperspectral Image by Slow Feature Analysis and Simplex Volume Ratio Approach. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 560–563. https://doi.org/10.1109/IGARSS.2019.8898127

Du, B., Wang, Y., Wu, C., & Zhang, L. (2018). Unsupervised Scene Change Detection via Latent Dirichlet Allocation and Multivariate Alteration Detection. IEEE Journal of Selected Topics in *Applied Earth Observations and Remote Sensing, 11*(12), 4676–4689. https://doi.org/10.1109/JSTARS.2018.2869549

Du, W., Wang, Y., & Qiao, Y. (2018). Recurrent Spatial-Temporal Attention Network for Action Recognition in Videos. *IEEE Transactions on Image Processing, 27*(3), 1347–1360. https://doi.org/10.1109/TIP.2017.2778563

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).

E-ISSN: 2948-5231

Gao, M., Zhou, X., Liu, Q., Yang, G., & Wu, C. (2022). Automatic monitoring of farmland occupation by farmhouse based on deep learning network. *Bulletin of Surveying and Mapping, 3*, 47–53. https://doi.org/10.13474/j.cnki.11-2246.2022.0076

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hu, B., Zhang, W., Ma, T., & Zhao, Z. (2021). Optimization and Simulation of Farmland Protection Dynamic Monitoring System Based on Internet of Things Technology. *Wireless Communications and Mobile Computing*, 2021, 1–11. https://doi.org/10.1155/2021/6902998

Jabari, S., Rezaee, M., Fathollahi, F., & Zhang, Y. (2019). Multispectral change detection using multivariate Kullback-Leibler distance. *ISPRS Journal of Photogrammetry and Remote Sensing, 147*, 163–177. https://doi.org/10.1016/j.isprsjprs.2018.11.014

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment, 132*, 159–175. https://doi.org/10.1016/j.rse.2013.01.012

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017). Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment. *International Journal of Applied Earth Observation and Geoinformation, 54*, 15–27. https://doi.org/10.1016/j.jag.2016.08.010

Shang, C., Yang, F., Huang, B., & Huang, D. (2018). Recursive Slow Feature Analysis for Adaptive Monitoring of Industrial Processes. *IEEE Transactions on Industrial Electronics, 65*(11), 8895–8905. https://doi.org/10.1109/TIE.2018.2811358

Sicong Liu, Bruzzone, L., Bovolo, F., & Peijun Du. (2015). Hierarchical Unsupervised Change Detection in Multitemporal Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing, 53*(1), 244–260. https://doi.org/10.1109/TGRS.2014.2321277

Wu, Z., Shen, C., & van den Hengel, A. (2019). Wider or Deeper: Revisiting the ResNet Model for Visual Recognition. *Pattern Recognition, 90*, 119–133. https://doi.org/10.1016/j.patcog.2019.01.006

Zhang, C., Yue, P., Tapete, D., Jiang, L., Shangguan, B., Huang, L., & Liu, G. (2020). A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing, 166*, 183–200. https://doi.org/10.1016/j.isprsjprs.2020.06.003

This work is licensed by the Creative Commons Attribution 4.0 International (CC By 4.0) (http://creativecommons.org/licenses/by/4.0/).
E-ISSN: 2948-5231