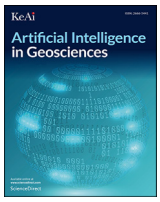
Artificial Intelligence in Geosciences 2 (2021) 26–34

Contents lists available at ScienceDirect

Artificial Intelligence in Geosciences

journal homepage: www.keaipublishing.com/en/journals/artificial-intelligence-in-geosciences



Capsule network-based approach for estimating grassland coverage using time series data from enhanced vegetation index

Yaqi Sun, Hailong Liu *, Zhengqiang Guo

School of Resources and Environment, University of Electronic Science and Technology of China, 611731, Chengdu, China

ARTICLE INFO

Keywords:
Classification
Deep learning
Grassland coverage
Remote sensing

ABSTRACT

The degradation and desertification of grasslands pose a daunting challenge to China's arid and semiarid areas owing to the increasing demand for them in light of the rise of animal husbandry. Monitoring grasslands by using big data has emerged as a popular area of research in recent years. As grassland degradation is a slow and gradual process, the accurate identification of grassland cover is key to monitoring it. Vegetation coverage is currently monitored mainly by combining inversion-based methods with field surveys, which requires significant human effort and other resources and is thus unsuitable for use at a large scale. We proposed to use time series from the enhanced vegetation index (EVI) in capsule network-based methods to identify grasslands. The process classified grassland coverage into four levels, high, medium, low, and other, based on Landsat images from 2019. The accuracy in classifying the grasslands at each level was higher than 90%, with an overall accuracy of 96.32% and a kappa coefficient of 0.9508. The proposed method outperformed the SVM, RF, and LSTM algorithms in terms of classification accuracy.

1. Introduction

The grassland is the largest terrestrial ecosystem in China, covering more than 40% of the country's territory. About 78% of this grassland is located in arid and semiarid regions (Ren et al., 2016), and it plays a critical role in facilitating animal husbandry and environmental conservation (Akiyama and Kawamura, 2007). Grassland degradation has been a worldwide issue in recent decades, and about 90% of the entire useable grassland in China has degraded to varying extents (Nan, 2005). The degraded area in China accounted for 22.7% of the total area of grassland from 1982 to 2010 (Zhou et al., 2017).

To quickly and accurately monitor the degradation in grassland, one uses the critical proxy of vegetation coverage to estimate essential parameters and functions of the ecosystem through remote sensing (Lehnert et al., 2015). Two approaches are widely used with satellite remote sensing datasets: spectral mixture analysis (SMA) and empirical models (Ge et al., 2018). The former decomposes each pixel in an image into a linear component of the reference spectrum, referred to as an endmember (Jiapaer et al., 2011). The model of pixel dichotomy is a simplified linear SMA to simulate vegetation cover rapidly. Empirical models represent the relationship between remote sensing data and data from field surveys. The remote sensing data of a band or a combination of bands have been used. Van et al. (2008) used Landsat ETM + data to estimate urban vegetation cover. They obtained accurate values for the regression of vegetation coverage as well as bands 2, 3, 5, and 7 of remote sensing images. Remote sensing data can also be converted into vegetation indices (Patel et al., 2007; North, 2002; Rundquist, 2002). Kergoat et al. (2015) showed that the ratio of the 1.6 μm and 2.1 μm bands acquired with a ground radiometer was well suited for retrieving the fraction of vegetation cover in crop fields, fallows, and grasslands.

Compared with empirical models, the output of the SMA can be less accurate (Liu et al., 2014). Therefore, Duan et al. (2019) integrated the vegetation index (VI) and a large amount of information from the SMA to improve the estimation accuracy of rice yields. The results showed that this method could better predict the yields than the VI alone. However, to further improve accuracy, empirical models require many samples from field surveys and thus are unsuitable for use at a large scale. Methods based on machine learning (Lehnert et al., 2015; Goetz et al., 2003) and artificial neural networks (Xie et al., 2009) can be alternatives.

Machine learning can be more accurate than the traditional regression model. However, their accuracy is affected by the modes of determining the required parameters and the choice of training samples (Huang et al., 2008). Ge et al. (2019) developed a gradient regression lifting-tree algorithm using hyperspectral image data acquired during

* Corresponding author.
E-mail address: liuhl@uestc.edu.cn (H. Liu).

https://doi.org/10.1016/j.aiig.2021.08.001
Received 14 April 2021; Received in revised form 2 August 2021; Accepted 4 August 2021
Available online 8 September 2021

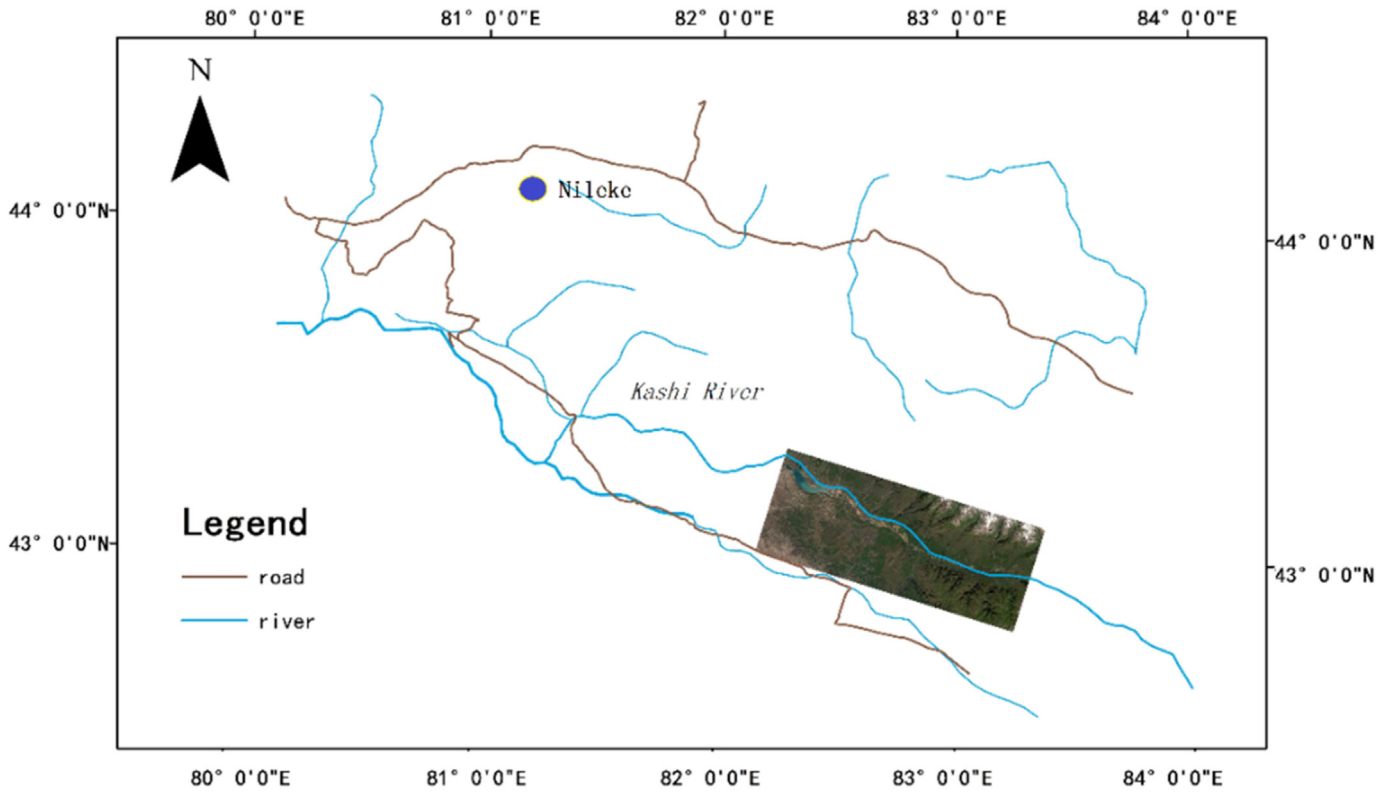


Fig. 1. The study area is located in the middle reach of Kashi River, Xinjiang.

flight. They used it to establish a model to estimate soil moisture content based on geographically weighted regression. This model has a high inversion accuracy.

Due to their impressive feature extraction and expression capabilities, deep learning algorithms are commonly employed to monitor changes in vegetation based on remote sensing images (Dinh et al., 2018). Marc et al. (2017) used a long short-term memory (LSTM) network to classify crops coupled with a time series of satellite images. The results showed that it was more accurate for crop classification than the convolutional neural network (CNN) and supper vector machine (SVM) algorithms. Han et al. (2015) proposed a framework to detect geospatial objects by combining a weakly supervised learning (WSL) approach with advanced feature learning. The experimental results with three optical datasets in benchmark tests verified its effectiveness compared with several state-of-the-art methods for object detection. Pan et al. (2018) proposed a multi-granularity scanning method, using the spectral relationship between different bands and the spatial correlation among adjacent pixels. In the context of a limited number of samples, they used a semi-supervised method to generate a convolution kernel. The technique was superior to some of the most advanced hue, saturation, and intensity (HSI) classification approaches on small datasets.

Although the deep learning approach generally has a higher accuracy of change assessment, it requires many samples for training. The capsule network method, a deep learning method proposed by a team led by Geoffrey Hinton (Sabour S. et al., 2017), originated in the context of the loss of location information in a neural network (Lecun et al., 1998). It overcomes the shortcoming of CNN algorithms whereby they are not sensitive to location information by adjusting their structure. The mechanism of the reception and identification of visual information by the human eye forms the archetype for the capsule network. The brain parses a hierarchical representation of the visual scene and matches the learned patterns with relationships stored by it (Sara G et al., 2019). Compared with other deep learning methods, it has a significant advantage in accurately identifying targets using a small amount of data. We used the capsule network method with a one-dimensional (1-D) convolution as the feature extraction layer to adequately extract features from time-series data.

One of the difficulties in monitoring grassland coverage through remote sensing is cloud cover and human-related factors. We used the time series of the enhanced vegetation index (EVI) to eliminate the cloud effects in a given remote sensing image. The EVI reflects vegetation growth and reduces interference by soil as background (Huete et al., 2002). It solves the problem whereby the normalized difference vegetation index (NDVI) is quickly saturated in high vegetation coverage and is easily affected by soil in low vegetation coverage (Wang et al., 2006). In addition, the EVI can compensate for the absorption of red light by residual aerosols (Rocha et al., 2009). Seasonal trend analysis (STA) was performed using an enhanced monthly EVI time series from 2005 to 2016 that was edited from MOD13Q1 products (MODIS Terra, https://lads web.modaps.eosdis.nasa.gov) used to detect changes in mangrove cover in Sinaloa state of northwestern Mexico (Berlanga-Robles and Ruiz-Luna, 2020). Despite the coarse resolution, the MODIS EVI products helped distinguish short- and long-term changes in mangrove cover, particularly in large ecosystems.

2. Materials and methods

2.1. Study area

The study area (latitude 43°34′–43°53′N, longitude 82°50′–83°55′E), located in the middle reaches of Kashi River, is part of Nileke County in Xinjiang, China. Xinjiang has the third largest area of grassland and one of the five largest pastoral areas in China (Zhang et al., 2013). The grassland guarantees animal husbandry and provides the last line of defense against desertification in the region.

The sites selected (Fig. 1) have a temperate and semiarid continental climate with an average annual temperature of 5.6 °C (Yi et al., 2006). Precipitation is plentiful but unevenly distributed in the region, with

Table 1
Main parameters of Landsat and GF-1 datasets.

Satellite	Sensor	Bands	Resolution (m)	Revisit time (days)	Data source
Landsat-7	ETM+	1-5, 7	30	16	USGS Earth Explorer
Landsat-8	OLI	1–7	30	16	https://earthexplorer.usgs.gov/
GF-1	PMS1, PMS2	5	2, 8	4	http://www.cresda.com/CN/

more in the eastern than western parts and more in the mountain than plain. Snow accumulates for 20–160 days every year, with a 40–50 cm thickness, and a maximum thickness can be around 90 cm on average in the early spring. Snowfall accounts for about one-third of the precipitation. The average annual evaporation is 1471.6 mm, increasing from west to east. The mild climate and sufficient water flow created abundant grassland.

Land use in the study area is dominated by grass land, accounting for 87.51% of the county's total area, followed by other land types accounting for 11.01%, and developed areas accounting for only 1.47%. The study area is a pastoral area with grassland and forest land dominance.

2.2. Data

Grassland in the study area grows from late April to late September. We downloaded 14 Landsat datasets from May to the end of September 2019, including Landsat-8 OLI and Landsat-7 ETM + datasets (Table 1). Sample data for classification were obtained from Landsat-8 OLI on August 11, 2019. The Landsat data were subjected to calibration and atmospheric correction using the fast line-of-sight atmospheric analysis of hypercubes (FLAASH) algorithm (Cooley et al., 2002). The algorithm is widely used for atmospheric correction owing to its high calibration accuracy and easy operation.

Two Chinese GF-1 (Table 1) optical images (Panchromatic band) were used for visual interpretation. A field study was conducted in July 2019 to verify the visual interpretation accuracy.

Fig. 2. Sample sites in the study area. Grasslands with low (a), medium (b), and high (c) coverages.

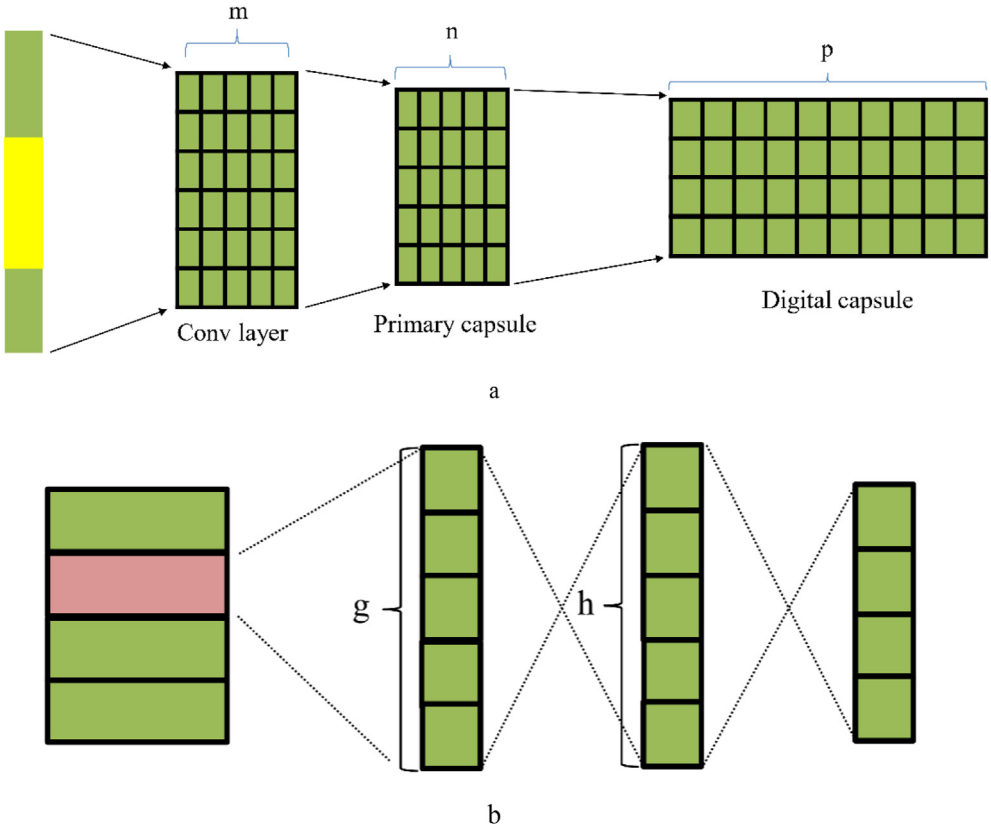


Fig. 3. (a) The improved capsule network, and (b) the reconstructed layer.

Thirty-one samples were collected around the center of the study site by using a global positioning system (GPS) unit and a tape measure. Over a selected sample area (30 m × 30 m), we recorded the range of grassland coverage by visual estimation, which was fast and adequate for our purposes (Fig. 2). In total, were recorded eleven high-, nine medium-, and four low-coverage grasslands, four areas of bare land, and three types of other land covers.

2.3. Methods

2.3.1. Improved capsule network algorithm

A capsule network is composed of convolutional, capsule, and reconstruction layers. It removes the original pooling layer of the CNN that causes the loss of location information. The capsule network uses convolutional layers to extract features, and the capsule layers are responsible for learning the features. The reconstruction layer regularizes the capsule network to prevent overfitting. The network has a different routing algorithm from that of a typical neural network. The routing algorithm determines data transmission between the capsule layers and the safety of all information and eliminates the defects of the CNN (Sara Ge et al., 2019). Three relevant parameters here are the similarity measure (b), coupling coefficient (c), and weight matrix (W). i and j are capsules in the current and subsequent capsule layers, respectively, and are associated with the routing algorithm. The coupling coefficient between capsules i and j is c_{ij}, and W_{ij} is the weight matrix between i and j. The similarity measure between i and j is b_{ij}. In every iteration, c_{ij} was updated by using the softmax function (Formula (1)), which ensured that the sum of the coupling coefficients between capsules was one (Sabour S. et al., 2017):

$$c_{ij} = soft\max(b_i) \tag{1}$$

We used W_{ij} with output u_i of the previous layer of capsule i calculating the weighted sum s_j according to c_{ij}. Capsule j was then given as

$$\widehat{u}_{j|i} = W_{ij} u_i \tag{2}$$

$$s_j = \sum_i c_{ij} \widehat{u}_{j|i} \tag{3}$$

To ensure that the length of the capsule was between zero and one and to introduce nonlinear factors, we used a squashing function to determine the final capsule of j, which was given as

$$v_j = \frac{||s_j||^2}{1 + ||s_j||^2} * \frac{s_j}{||s_j||} \tag{4}$$

b_{ij} were initialized to zero before applying the routing algorithm and updated with

$$b_{ij} = b_{ij} + \widehat{u}_{j|i} v_j \tag{5}$$

The weights and the similarity between capsules determine data transmission between the capsule layers. More similar capsules can help derive more information from the lower capsule layer that can be used to identify minor differences robustly.

It should be noted that the described capsule network is designed for data with multiple dimensions. According to the above characteristics of the capsule network, we propose a pixel-oriented approach to classification to fit the 1-D EVI data to extract features from the time series data. We change the feature extraction layer to a 1-D convolution layer, and the final model is composed of two convolutional layers, two capsule layers, and two reconstruction layers. In Fig. 3(a), m referred to the number of convolutional kernels, n to the number of second convolutional kernels, and p to the length of the digital capsule. Fig. 3(b) shows the reconstructed part of the capsule network. It is still a general neural network, with the input data indicated by the pink rectangle, which has the highest modular length. The letters g and h referred to the number of hidden units.

Table 2
Numbers of training and testing samples for each class.

Class	Training	Testing
Other	6400	1600
Low-coverage grassland	8000	2000
Medium-coverage grassland	8000	2000
High-coverage grassland	8000	2000

2.3.2. Model settings

2.3.2.1. Selecting the samples. To ensure accurate and efficient monitoring of grasslands at a large scale, we propose a method to classify grassland coverage into four levels according to China's system to classify land use and land cover (LULC) types based on remote sensing monitoring data (National Standard of the People's Republic of China, 2017). The four classes were: high-coverage grassland (≥50%), medium-coverage grassland (20%–50%), low-coverage grassland (5%–20%), and bare land (≤5%). Because other types of cover, including barren, water, and developed areas, occupied a small space in the study area, we eventually used the four grassland classes and others. We randomly selected representative samples and divided them into a training set and a test set with a ratio of 4:1 based on experience (Table 2).

2.3.2.2. Extracting EVI time series based on multi-temporal remote sensing images. The EVI, used to increase the sensitivity of vegetation signals in regions with high biomass (Huete et al., 2002), is

$$\mathrm{EVI} = G^* \frac{B_{NIR} - B_R}{B_{NIR} + C_1 * B_R - C_2 * B_B + L} \tag{6}$$

where B_{NIR}, B_R, and B_B are the atmospheric or partially corrected atmosphere surface reflectances in the near-infrared (NIR), red, and blue bands, respectively. L is the canopy background adjustment used to address nonlinear, differential-NIR, and red radiant transfer through the canopy. C_1 and C_2 are coefficients of aerosol resistance. The blue band is used to correct the influence of aerosol on the red band. The coefficients in (6) are set to $L = 1$, $C_1 = 6$, $C_2 = 7.5$, and G (gain factor) = 2.5 (Merchant, 2000).

A total of 14 Landsat images (Table 1) were used. The cloud covers were less than 7% in five images, 10%–20% in four, 35%-50%was in two, and higher than 50% in two. The Landsat-7 ETM + images had an additional gap of about 15% due to the failure of the scan line corrector. Areas with cloud cover were mainly distributed in mountainous areas and affected the identification of low-coverage grassland and bare land. To eliminate their effects in the EVI time series, we used the quality band of the remote sensing images to identify pixels representing cloud cover. These pixels and gap pixels were linearly interpolated using the nearest valid values before or after the time step (Zhong et al., 2019). The result is shown in Fig. 4. Landsat-7 data were at intervals 1, 3, …, and 13, but Landsat-8 datasets at intervals 2, 4, …, and 14.

The EVI time series were extracted from both Landsat 7 ETM+ and Landsat 8 OLI ing images. Given that each Landsat 7 ETM + remote image had a gap of up to 15%, we constructed two EVI time series for the sake of comparison. We used the method mentioned in the previous section for processing to obtain the first EVI time series. The second time series retained only the EVI of the corresponding Landsat 8 images in the first time series. We were able to eliminate the influence of some clouds in the Landsat 8 remote sensing images. However, the disadvantage of the second EVI time series was that its interval was long, up to 16 days, whereas the interval of the first EVI time series was eight days. We trained two classifiers of the capsule network over 50 rounds using the same parameter settings (Table 3).

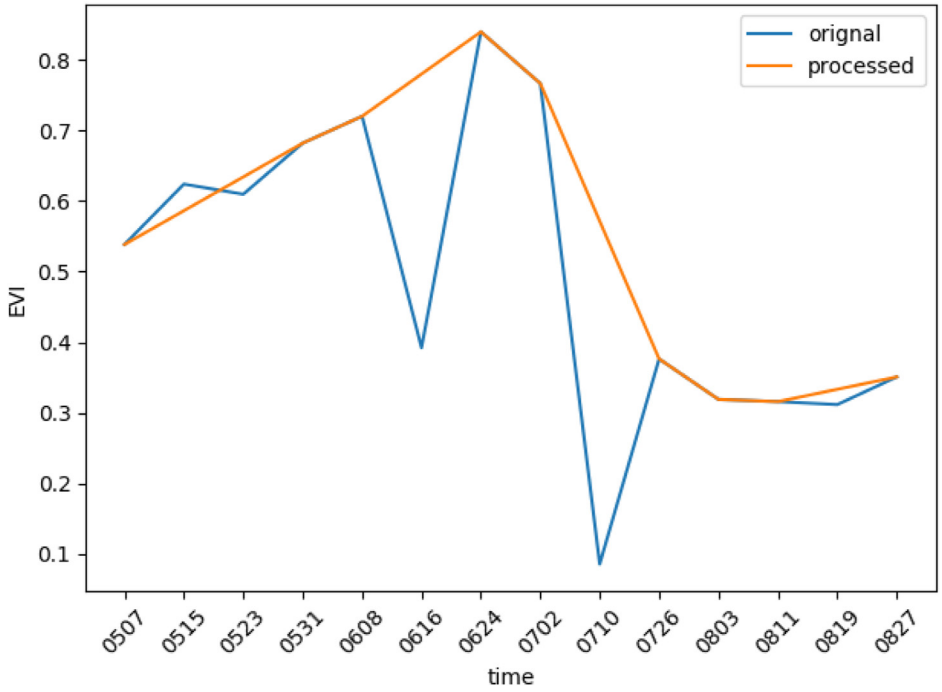


Fig. 4. Example of results of processing the time series.

Table 3
Parameters of the capsule network.

Parameter	m	n	p	G	h
setting	10	20	16	20	40

2.3.3. Algorithms for comparison

The SVM algorithm can find an optimal hyperplane using structural risk minimization. It ensures that the distance between data samples on both sides of the plane and the hyperplane is maximized (Hoteling H et al., 1993). Burai et al. (2015) used the algorithm to identify 20 kinds of alkaline grassland vegetation with an overall accuracy of classification of 82.60%. They found that the algorithm was not sensitive to the number of training samples and was suitable for classification when the number of training samples is limited.

The random forest (RF) model consist of decision trees. Each tree is a classifier, and the RF model integrates multiple decision trees. The integration of multiple classifiers improved accuracy (Verikas et al., 2011). Thes model has been applied to make spatial predictions of fractional wood coverage in the Molopo area by training a random forest algorithm with the forward feature variable selection. Spatial cross-validation revealed a prediction error of 12% (Ludwig et al., 2019).

The long short-term memory (LSTM) neural network is a particular recurrent neural network (RNN) (Zazo et al., 2016). It can eliminate unnecessary information, retain valuable information for prediction, and thus deal with long-term dependence. The LSTM NN is widely used in time series processing because its unique structure helps identify the characteristics of time series samples.

3. Results

3.1. Classification of grassland coverage based on capsule network

We used the improved capsule network to classify grassland coverage in the Kashi River basin in 2019 (Fig. 5). Areas with low, medium, and high grassland coverages are shown in the light, lime, and forest green colors. Other land cover types are in black. The low-coverage grassland

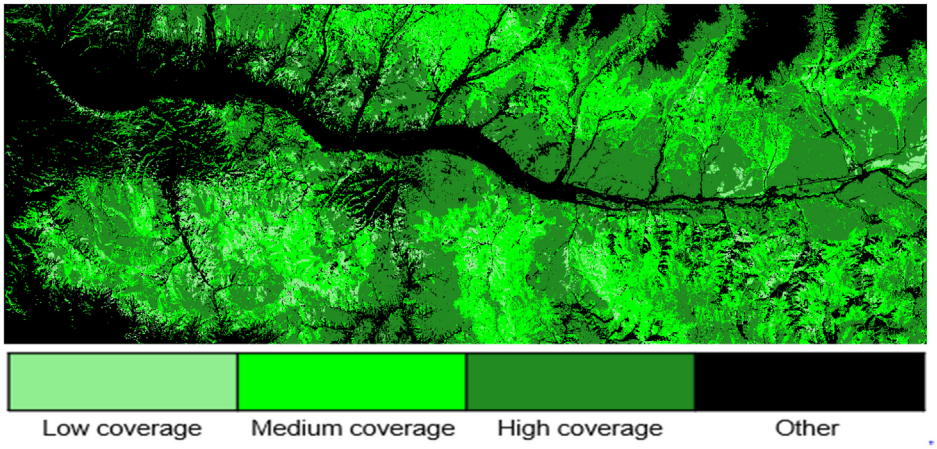


Fig. 5. Classification map of grassland coverage in Kashi River Basin in 2019.

Table 4
Kappa coefficients (%) of the capsule network approach.

Class	Capsule network
Other	96.31
Low coverage grassland	92.20
Medium coverage grassland	98.30
High coverage grassland	98.45
Overall	96.32
Kappa	95.08

was mainly distributed in the western part of the study area, and the medium-coverage grassland in the mountains and the north. The vast majority of high-coverage grassland was around on both sides of the river.

Table 4 shows that the proposed method records a good extraction accuracy on the four types of grassland coverage. It is 90% or higher for each class, with an overall accuracy of 96.32%. The kappa coefficient was 0.9508. Except for its performance on the low-coverage grassland (92.2%), the proposed method had an accuracy higher than 95% in identifying the other types of coverage. The accuracy in delineating the high-coverage grassland was the highest, at 98.45%.

To compare the classification accuracy of the proposed method with observations of the samples, we combined the map of vegetation coverage with that of the sampling points, obtaining an overall accuracy of 90.32% (Fig. 6). Of the 31 samples, the vegetation coverage of only three was inconsistent with the results of classification obtained by the capsule network.

3.2. *Effect of EVI time series on the accuracy of classification*

The capsule network used the marginal loss as an indicator of the effects of training. We used the same loss function in this study. The curves of the loss function (Fig. 7) of both classifiers show that the training yielded the best results.

As shown in Table 5, the effect of classification of classifier I was generally better than that of classifier II for all types of land, but especially for other land cover types. Although the Landsat 7 data occupied a relatively large proportion, the EVI time series with a short interval generated better results.

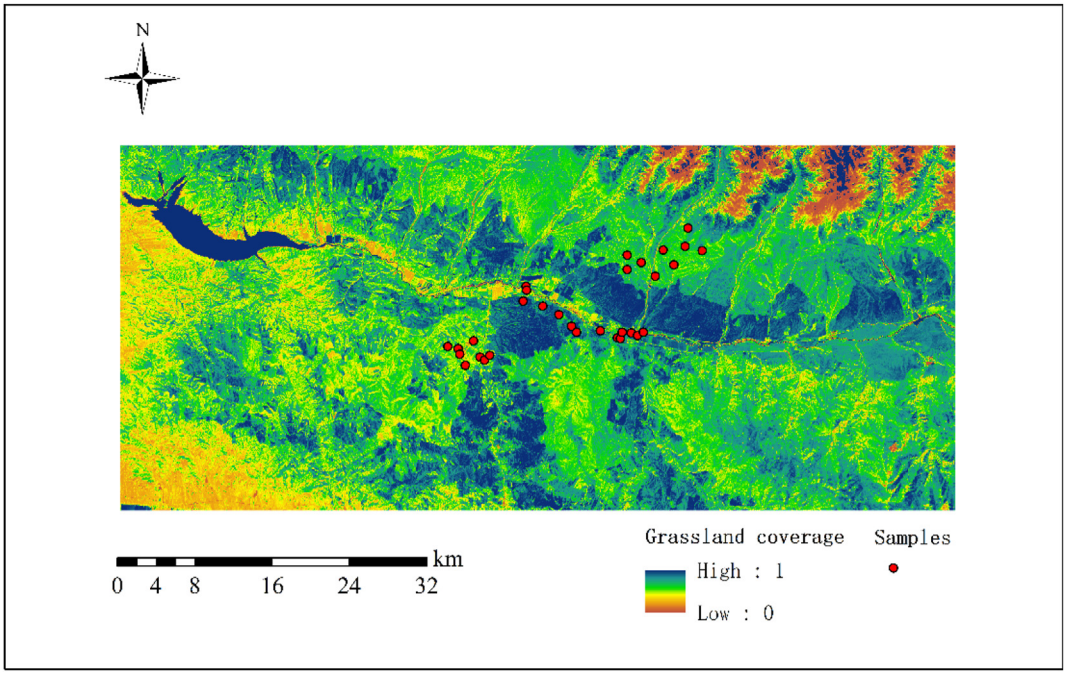


Fig. 6. The accuracy evaluation of the proposed method.

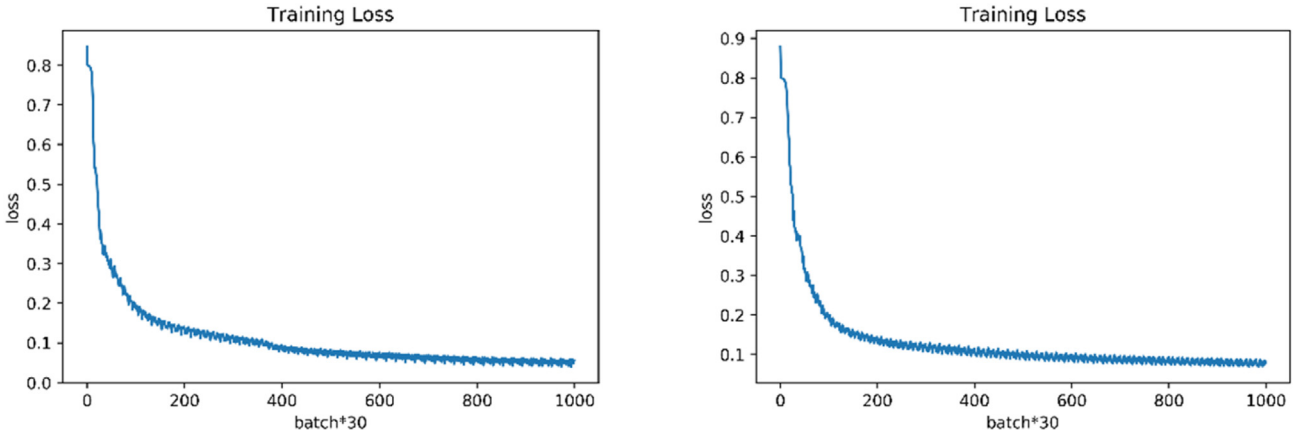


Fig. 7. Curves of the loss function of classifier I (left) and classifier II (right).

Table 5
Kappa coefficients (%) of the capsule network approach with two EVI time series.

Class	Classifier I	Classifier II
Other	96.31	77.75
Low coverage grassland	92.20	93.55
Medium coverage grassland	98.30	97.05
High coverage grassland	98.45	98.50
Overall	96.32	92.45
Kappa	95.08	89.87

Table 6
Results of incorrect classification of grassland types with various cloud coverages.

Affected EVI data	3	4	5	6	7
Misclassified	1	13	25	74	80
Total	155	375	780	1610	1922
Ratio (%)	0.645	3.47	3.21	4.60	4.16
Affected EVI data	8	9	10	11	
Misclassified	59	24	4	0	
Total	1552	973	227	6	
Ratio (%)	3.80	2.47	1.76	0	

Table 7
Kappa coefficients (%) of four methods.

Class	Capsule network	SVM	RF	LSTM
Other	96.31	88.31	90.43	78.06
Low coverage grassland	92.20	78.65	89.95	90.35
Medium coverage grassland	98.30	96.7	97.05	96.95
High coverage grassland	98.45	97.9	98.40	98.75
Overall	96.32	90.50	94.14	91.71
Kappa	95.08	87.32	92.17	88.89

Values of the minimum, maximum, mean, and standard deviation of cloud cover in the EVI time series were 0.09%, 100%, 30.58%, and 34.7%. To understand the impact of the proportion of cloud cover in the EVI time series on classification, we analyzed the response relationship of cloud cover with the ratio of incorrect classification (Table 6). The results showed no significant impact (Landsat 7) on the classification results. However, all 11 phases of the time series data affected were sporadic, and six pixels were not representative of the study area. When three phases of data in the EVI time series were impacted, the error rate was close to zero. When more phases were affected, the error rate was between 1% and 5%. Thus, remote sensing images with less cloud cover were more suitable for identifying vegetation cover.

4. Discussion

To verify the accuracy of the proposed capsule network in terms of estimating grassland coverage, we conducted a comparative experiment (Table 7) involving the SVM, RF, and LSTM algorithms.

Compared with the other three methods, the proposed method improved the extraction accuracy of four different coverage grasslands. The accuracy of grassland coverage monitoring and the overall accuracy and kappa coefficient reached 96.32% and 95.08%. Although the accuracy of extraction of the SVM algorithm for bare land was higher than that of the LSTM NN, its accuracy in identifying low-coverage grassland was only 78.65%. Its overall classification accuracy was lower than those of the LSTM and RF methods. The RF algorithm had the highest classification accuracy of the three compared methods, but its accuracy of identifying low-coverage grassland was less than 90%. While the LSTM NN is widely used for time series processing, it yielded less than 80% classification accuracy. Only on high-coverage grassland did it perform well, with an accuracy of 98.75%. Its overall classification accuracy was only slightly higher than that of the SVM algorithm.

To analyze the accuracy of the capsule network in extracting spatial information on grass coverage, we compared its results with those of the SVM, RF, and LSTM algorithms, as shown in Fig. 8. These four methods have no significant differences in the other category but are mainly reflected in the classification of grassland coverage. To compare the capsule network approach and the three algorithms, we subtracted the resulting map of the former from that of each of the latter. The consistency of the classification accuracy between the capsule network method and each method is given in Table 8. The classification results of the capsule network matched those of the SVM, RF, and LSTM algorithms at 84.66%, 85.03%, and 84.49%, respectively. There were very few cases in which the results did not match. In most of them, the capsule network had identified a given area as high coverage, whereas the other networks had classed it as medium coverage or vice versa. This is because high-coverage and medium-coverage grasslands are challenging to distinguish. Because of the complexity of samples considered and the inherent advantages of the deep learning used in the capsule network, the proposed method was more accurate on all four coverage types than the other methods and thus had a higher overall accuracy. It was.

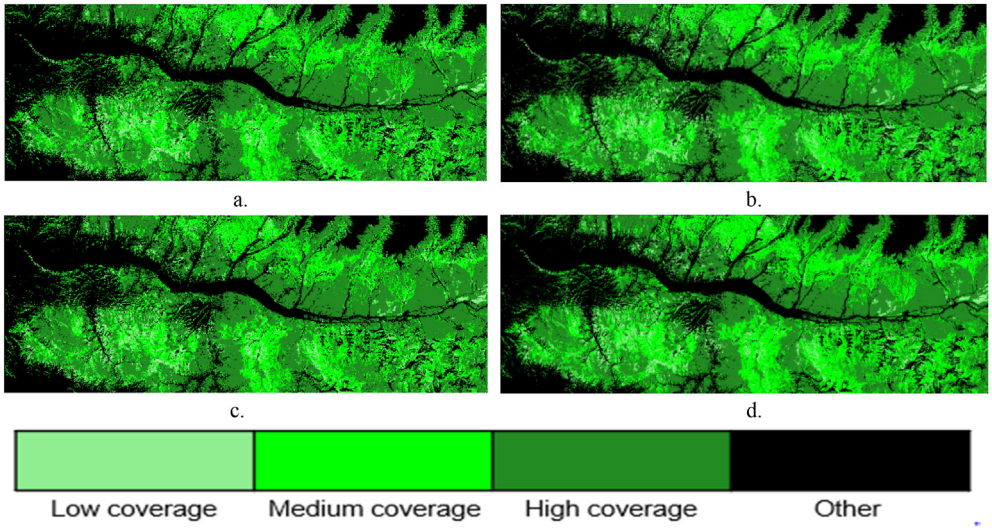


Fig. 8. Classification maps of grassland coverage, Kashi River Basin in 2019 after SVM (a), RF (b), LSTM (c), and capsule network (d) algrithms.

Table 8
Inter-comparision of classified outputs.

		SVM				
		High coverage	Medium coverage	Low coverage	Other	Consistent result
Capsule network	High coverage	16.4808	6.4555	2.3812	0	
	Medium coverage	3.3551	29.7992	0.7277	0	
	Low coverage	1.698	0.7258	2.9201	0	
	Other	0	0	0	35.4566	
	Consistent result					84.6567
		RF				
		High coverage	Medium coverage	Low coverage	Other	Consistent result
Capsule network	High coverage	17.0778	5.7727	2.467	0	
	Medium coverage	3.4907	29.5443	0.8469	0	
	Low coverage	1.699	0.6953	2.9497	0	
	Other	0	0	0	35.4566	
	Consistent result					85.0284
		LSTM				
		High coverage	Medium coverage	Low coverage	Other	Consistent result
Capsule network	High coverage	15.6737	7.3141	2.3298	0	
	Medium coverage	2.9575	30.1949	0.7295	0	
	Low coverage	1.5147	0.6677	3.1615	0	
	Other	0	0	0	35.4566	
	Consistent result					84.4867

5. Conclusions

This study proposed a quick and efficient method to classify grassland coverage that can classify grasslands into four quantitative types: high, medium, low, and other. EVI time series data of the cycle of grass growth were used to predict grassland coverage. We used the capsule network to classify remote sensing images. The proposed method yielded an accuracy of classification higher than 90% on each grassland class, with an overall accuracy of 96.32%. Its kappa coefficient was 0.9508. It also outperformed the SVM, RF, and LSTM algorithms in accuracy assessment in the comparative analyses.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the International Science and Technology Cooperation Project of Sichuan (2020YFH0067) and the China Academy of Sciences Strategic Leading Science and Technology Project (XDA20060303).

References

Akiyama, T., Kawamura, K., 2007. Grassland degradation in China: methods of monitoring, management and restoration. Grassl. Sci. 53 (1), 1–17.
Berlanga-Robles, C.A., Ruiz-Luna, A., 2020. Assessing seasonal and long-term mangrove canopy variations in Sinaloa, northwest Mexico, based on time series of enhanced vegetation index (EVI) data. Wetl. Ecol. Manag. 28 (2), 229–249.
Burai, P., Deák, B., Valkó, O., Tomor, T., 2015. Classification of herbaceous vegetation using airborne hyperspectral imagery. Rem. Sens. 7 (2), 2046–2066.
Cooley, T., Anderson, G.P., Felde, et al., 2002. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. IEEE International Geoscience and Remote Sensing Symposium 3, 1414–1418.
Dinh, H., Gaetano, D., Lalande, N., et al., 2018. Deep recurrent neural networks for winter vegetation quality mapping via multi-temporal SAR sentinel-1. Geosci. Rem. Sens. Lett. IEEE 15 (3), 464–468.
Duan, B., Fang, S., Zhu, R., et al., 2019. Remote estimation of rice yield with unmanned aerial vehicle (UAV) data and spectral mixture analysis. Front. Plant Sci. 10, 204.
Ge, J., Meng, B., Liang, T., et al., 2018. Modeling alpine grassland cover based on MODIS data and support vector machine regression in the headwater region of the Huanghe River, China. Remote Sens. Environ. 218, 162–173.
Ge, X., Wang, J., Wang, J., Ding, J., et al., 2019. Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring. PeerJ, Preprints 7, e27630v1.
Goetz, S.J., Wright, R.K., Smith, A.J., et al., 2003. IKONOS imagery for resource management: tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region. Rem. Sens. Environ. 88 (1), 195–208.
Huete, A., Didan, K., Miura, T., et al., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Rem. Sens. Environ. 83 (1), 195–213.
Huang, C., Song, K., Kim, S., et al., 2008. Use of a dark object concept and support vector machines to automate forest cover change analysis. Rem. Sens. Environ. 112 (3), 970–985.
Han, J., Zhang, D., Cheng, G., et al., 2015. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. IEEE Trans. Geosci. Rem. Sens. 53 (6), 3325–3337.
Hoteling, H., 1993. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24 (6), 417–520.
Jiapaer, G., Chen, X., Bao, A., 2011. A comparison of methods for estimating fractional vegetation cover in arid regions. Agric. For. Meteorol. 151 (12), 1698–1710.
Kergoat, L., Hiernaux, P., Dardel, et al., 2015. Dry-season vegetation mass and cover fraction from SWIR1.6 and SWIR2.1 band ratio: ground-radiometer and MODIS data in the Sahel. ITC Journal 39 (10), 56–64.
Lehnert, L.W., Meyer, H., Wang, Y., et al., 2015. Retrieval of grassland plant coverage on the Tibetan Plateau based on a multi-scale, multi-sensor and multi-method approach. Rem. Sens. Environ. 164, 197–207.
Liu, B., Shen, W., Lin, N., et al., 2014. Deriving vegetation fraction information for the alpine grassland on the Tibetan plateau using in situ spectral data. J. Appl. Remote Sens. 8 (1), 083630.
Lecun, Y., Bottou, L., Bengio, Y., et al., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.
Ludwig, M., Morgenthal, T., Detsch, F., et al., 2019. Machine learning and multi-sensor based modelling of woody vegetation in the Molopo Area, South Africa. Rem. Sens. Environ. 222, 195–203.
Merchant, J.W., 2000. Remote sensing of the environment: an Earth resource perspective. Cartogr. Geogr. Inf. Sci. 27 (4), 311.
Nan, Z., 2005. The grassland farming system and sustainable agricultural development in China. Grassl. Sci. 51 (1), 15–19.
National Standard of the People's Republic of China, 2017. Current Land Use Classification: GB/T 21010-2017[s]. Standards Press of China, Beijing.
North, P.R.J., 2002. Estimation of fAPAR, LAI, and vegetation fractional cover from ATSR-2 imagery. Rem. Sens. Environ. 80 (1), 114–121.
Patel, N.K., Saxena, R.K., Shiwalkar, A., 2007. Study of fractional vegetation cover using high spectral resolution data. Journal of the Indian Society of Remote Sensing 35 (1), 73–79.
Pan, B., Shi, Z., Xu, X., 2018. MugNet: deep learning for hyperspectral image classification using limited samples. ISPRS J. Photogrammetry Remote Sens. 145, 108–119.
Ren, Y., Lü, Y., Fu, B., 2016. Quantifying the impacts of grassland restoration on biodiversity and ecosystem services in China: a meta-analysis. Ecol. Eng. 95, 542–550.

Rundquist, B.C., 2002. The influence of canopy green vegetation fraction on spectral measurements over native tallgrass prairie. Rem. Sens. Environ. 81 (1), 129–135.
Rocha, A.V., Shaver, G.R., 2009. Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. Agric. For. Meteorol. 149 (9), 1560–1563.
Sabour, S., Frosst, N., Hinton, G., 2017. Dynamic routing between capsules. In: 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA.
Sara, G., Maria, T.V., Marta, S., et al., 2019. Unsupervised neural network for homography estimation in capsule endoscopy frames. Procedia Computer Science 164, 602–609.
Verikas, A., Gelzinis, A., Bacauskiene, M., 2011. Mining data with random forests: a survey and results of new tests. Pattern Recogn. 44 (2), 330–349.
Wang, Z., Liu, C., Chen, W., 2006. Preliminary comparison of MODIS-NDVI and MODIS-EVI in eastern Asia. Geomatics Inf. Sci. Wuhan Univ. 31 (5), 407–410.
Xie, Y., Sha, Z., Yu, M., et al., 2009. A comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China. Ecol. Model. 220 (15), 1810–1818.
Yi, L., Li, J., 2006. Discussion on optimum design of Plant landscape in Tangbula National forest Park. Forestry of Xinjiang (4), 35–37 (in Chinese).
Zhang, X., Liao, C., Li, J., et al., 2013. Fractional vegetation cover estimation in arid and semiarid environments using HJ-1 satellite hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 21, 506–512.
Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. Rem. Sens. Environ. 221, 430–443.
Zhou, W., Yang, H., Huang, L., et al., 2017. Grassland degradation remote sensing monitoring and driving factors quantitative assessment in China from 1982 to 2010. Ecol. Indicat. 83, 303–313.
Zazo, R., Lozano-diez, A., Gonzalez-Dominguez, J., et al., 2016. Language identification in short utterances using long short-term memory (LSTM) Recurrent Neural Networks. PloS One 11 (1), 1–17.