Research on classification of nutrient forage species based on VGG-16 network

Jvsong Huang[*1], Jiandong Fang[*2], Yvdong Zhao[*3], Bajin Li[*4]

(*1. Inner Mongolia University of Technology, Inner Mongolia Key Laboratory of Perceptive Technology and Intelligent System, Hohhot, Inner Mongolia, 010080, China, E-mail: 1837629094@qq.com

*2. Inner Mongolia University of Technology, Inner Mongolia Key Laboratory of Perceptive Technology and Intelligent System, Hohhot, Inner Mongolia, 010080, China, E-mail: fangjd@imut.edu.cn,

*3. Inner Mongolia Key Laboratory of Perceptive Technology and Intelligent System, Hohhot, Inner Mongolia, 010080, China, E-mail: zhaoyvdong@163.com

*4. Inner Mongolia University of Technology, Inner Mongolia Key Laboratory of Perceptive Technology and Intelligent System, Hohhot, Inner Mongolia, 010080, China, E-mail: libajin@imut,edu.cn)

The Project Supported by: Natural Science Foundation of Inner Mongolia (2019MS06023)；Tackling Key Scientific and Technological of Inner Mongolia (2019GG334)；

Abstract: Aiming at the problem of time-consuming, labor-intensive and inaccurate forage recognition traditionally relying mainly on manual direct observation and manual feature extraction, a nutritional forage classification model based on improved VGG-16 convolutional neural network is proposed. Based on the VGG-16 network model, the model optimizes the number of fully connected layers, and replaces the SoftMax classifier in the original VGG-16 network with a 5-label SoftMax classifier, optimizes the model structure and parameters, and shares the pre-trained model through migration learning The weight parameters of the middle convolutional layer and the pooling layer. Five types of nutrient forage images were collected from the forage experimental field in Siziwang Banner as the training data and test data of the classification model. The experimental results show that the model can accurately classify the five types of forage. In terms of the average recognition accuracy rate, the accuracy rate reached 91.3%, realizing the accurate classification of nutrient forages.

Keywords: VGG-16 network, Image classification, Transfer learning

1. INTRODUCTION

The grassland area in my country is now nearly 4 million square kilometers, and the number of grassland types and pasture species ranks first in the world. As an important base for the country's main animal husbandry development, the Inner Mongolia Autonomous Region has an impact on the survival and development of local people in pastoral areas due to the rational use and management of grassland resources, and is inseparable from the economic benefits, ecological balance and sustainable development of the region. As an important part of grassland resources and the basis of herbivorous livestock production, pasture is of great significance to the development of animal husbandry, environmental protection and soil conservation [1].

The classification and recognition of pasture images is one of the important basic requirements for the digital management of grassland resources. Fast and accurate classification and recognition is crucial to the planting and production of pasture. The past forage identification methods are mostly scientific researches showing that the application of artificial intelligence for the acquisition and classification of grassland resources related parameters will greatly improve the automation of forage production management and the real-time performance of related data acquisition [2]. Image classification and recognition can be roughly divided into two aspects based on hyperspectral images and digital images [3]. Due to the high cost and long time spent in hyperspectral image acquisition, it cannot be compared with digital images in terms of popularity and interpretability. Therefore, there are still certain limitations in the research of image classification and recognition. Compared with hyperspectral images, digital image acquisition is efficient, convenient, and low-cost, and the acquisition equipment is highly popular, and it is more widely used in practical applications. It can obtain image-related information from close distances and small scales, so it is feasible. In the hyperspectral image [4,5]. In 2010, Wang Jingxuan [12] and others used image processing technology to classify and recognize 14 legume pastures based on appearance features. By comparing features such as rectangularity, circularity, and horizontal and vertical axis ratios, the neural network (PNN) and Probabilistic Neural Network (BPN) were carried out respectively under classification experiments, the results were 85% and 82.4%. In 2016, Han Ding [13] and others used intelligent navigation vehicles to collect original grassland images, extracted RGB and HSV color moment features from two types of forage images of Leymus chinensis and Chenopodium chinense, and used the constructed 3-layer BP neural network for classification and recognition. , The final result is 89.5%.

In 2018, Han Ding [14] and others used the color and shape characteristics of four types of forage images as the basis for classification decision-making, based on the 3-layer BP neural network and PCA principal component analysis method for testing, and the final result was 85.5%. Through the above research, it can be seen that the existing pasture recognition methods are relatively small, with large errors, low recognition rates, cumbersome image acquisition methods, poor operability, and low automation, and the methods are mostly based on traditional machine learning methods. The feature information on the basis of classification decision is mostly derived from artificial feature design, and less use of deep learning-related ideas and techniques. Convolutional neural network [6], as one of the representative algorithms in the field of artificial intelligence, has powerful representation learning capabilities. The internal convolution kernel can directly extract deep features for classification decision-making from color images, without additional artificial feature engineering design for experimental data, and the VGGNet network [7] is a typical network in convolutional neural network classification and recognition. The training in the ImageNet database of 1 million images has powerful representation learning capabilities.

In summary, this article uses VGG-16 network and computer vision and other related technologies to propose a nutritional forage classification model based on the improved VGG-16 convolutional neural network, which can accurately classify five types of nutritional forages. In terms of the average recognition accuracy rate, the accuracy rate reached 91.3%, which is significantly better than the K-nearest neighbor algorithm (K-NN), support vector machine algorithm (SVM) and BP neural network algorithm. It can accurately realize the classification and recognition of nutrition forage images and improve The efficiency and automation level of pasture identification.

2. EXPERIMENTAL DATA AND ENVIRONMENT

2.1. Experimental data

The experimental research area is located in the Institute of Aerospace Information of the Chinese Academy of Sciences in Siziwang Banner, Inner Mongolia. The location coordinates are 111° 90′ east longitude, 41° 78′ north latitude, and 1000m above sea level. Located in the mid-temperate continental monsoon climate, the average precipitation is between 110-350mm, the annual average temperature is between 1-6°C, and the average annual difference is between 34-37°C.

The sample images selected in the experiment are all collected from the field under the natural background, which lays the foundation for the actual promotion and application in the later period. A total of 5 types of nutrition forage images were selected in the experiment, including 469 alfalfa, 329 Mongolian wheatgrass, 569 Nongke No. 1 tree Kochia, 305 rib vetch, and 317 corn. Each category randomly selects 90% as the training set and 10% as the validation set. The specific distribution is shown in Table 1.

Table. 1 Data set specific distribution table

project	Alfalfa	Agropyron mongolia	Kochia scoparia	Rib Vicia	corn
Training set	405	279	495	257	317
Validation set	44	30	54	28	35
Test set	20	20	20	20	20
total	469	329	569	305	372

2.2. Lab environment

The training equipment used in the experiment is a computer with GPU, the experiment platform is implemented based on the Pytorch deep learning framework, and the integrated development environment used is PyCharm. The specific hardware and software configuration of the training equipment used in the experiment is shown in Table 2.

Table. 2 Experimental platform configuration parameters

name	Configuration parameter
operating system	Windows10
RAM	64GB (Kingston DDR4 2666MHz)
CPU	Intel(R) i7-9700 CPU @ 3.00GHz
GPU	NVIDA GeForce RTX 2080Ti, 11GB
Integrated Development Environment	PyCharm 2020.2.1
GPU acceleration library	CUDA 10.1, CUDNN 8.0.3
Deep learning framework	Pytorch 1.6.0

3. RESEARCH METHODS

3.1. VGG-16 basic network

The CNN structure is shown in Figure 1, which includes a convolutional layer, a down-sampling layer, and a fully connected layer. The VGG-16 convolutional neural network contains 16 weight layers, which are composed of 13 convolutional layers and 3 fully connected layers [8]. Input a 224×224×3 image, adopt 3×3 filters in the convolution layer, and stack every 2 or 3 filters to form a convolution sequence, imitating the effect of a larger receptive field, sliding step size It is 1, and uses boundary filling to keep the front and back data dimensions unchanged; in the pooling layer, a 2×2 pooling window with a step size of 2 is used to reduce the size of the feature image after convolution and to ensure the translation of the model Invariance; in the fully-connected layer, it is composed of 3 continuous fully-connected channels, with the number of channels being 4096, 4096, and 1000 respectively; finally, the SoftMax classifier with 1,000 labels performs classification and output. The convolutional layer structure and parameters are shown in Table 3.

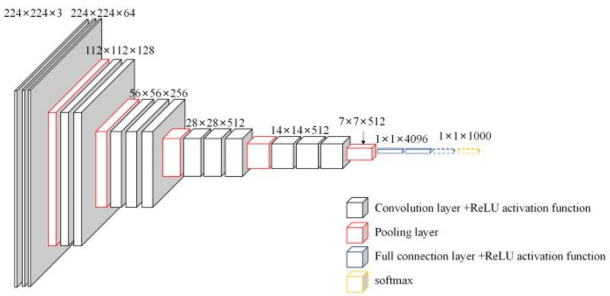


Fig.1 VGG-16 network structure diagram

Table 3 VGG-16 convolutional layer structure

Network layer (category)	Output type	Parameter
INPUT layer	(224, 224, 3)	0
Conv1 layer	(224, 224, 64)	1792
Conv2 layer	(224, 224, 64)	36928
Maxpool1 layer	(112, 112, 64)	0
Conv3 layer	(112, 112, 128)	73856
Conv4 layer	(112, 112, 128)	147584
Maxpool2 layer	(56, 56, 128)	0
Conv5 layer	(56, 56, 256)	295168
Conv6 layer	(56, 56, 256)	590080
Conv7 layer	(56, 56, 256)	590080
Maxpool3 layer	(28, 28, 256)	0
Conv8 layer	(28, 28, 512)	1180160
Conv9 layer	(28, 28, 512)	2359808
Conv10 layer	(28, 28, 512)	2359808
Maxpool4 layer	(14, 14, 512)	0
Conv11 layer	(14, 14, 512)	2359808
Conv11 layer	(14, 14, 512)	2359808
Conv11 layer	(14, 14, 512)	2359808
Maxpool5 layer	(7, 7, 512)	0
Convolution kernel parameter amount: 14717688		

3.2. Classification model of nutritional forage based on improved VGG-16 network

Based on the VGG-16 network structure and convolutional neural network, combined with the types and characteristics of nutrient forages, five types of nutrient forage classification models are constructed, as shown in Figure 2. The nutrient forage classification model is divided into four levels: convolutional layer, pooling layer, fully connected layer, SoftMax classification layer, among which the SoftMax classifier in the original VGG-16 network is replaced with the 5-label SoftMax classification layer. The activation function is the ReLU function, and the Adam optimization algorithm is selected.

The VGG-16 network is trained in the ImageNet database of 1 million images and has strong deep feature learning capabilities. It has a large number of parameters and weights that have been trained, especially the curves, edges, and edges of the image in the convolutional layer. The feature extraction of contour has strong ability [10]. In order to avoid training the entire network from scratch, reduce network training time and improve network training efficiency, the trained VGG-16 network can be used as the pre-training model of this model, and the pre-trained model parameters can be "migrated" to this research model. The parameters of the pre-training model optimize the model parameters of the convolutional layer to solve the problem of classification and identification of nutritional forages. The number of weight parameters of VGG-16 is 1.4×107, of which 3 fully connected layer parameters have a high degree of concentration. The parameters of VGG-16 are designed for 1000 classification categories, while this research only focuses on the classification of 5 categories. . Therefore, it is proposed to replace the original 3 fully connected layers with 2 fully connected layers, the first fully connected layer has 4096, and the second fully connected layer has 5, in order to improve the recognition accuracy and efficiency of the model. The improved model structure is shown in Figure 3.

The main operation process of model training is as follows:

- Enter 5 kinds of nutrient forage samples. Randomly extract 90% of the 5 kinds of nutrition forage pictures from the nutrition forage image database, and input them as the training sample set.
- Pretreatment. Filter the forage image, delete the fuzzy image, and standardize the input image to the same resolution size.
- Construct a recognition model for nutrient forages. Based on the VGG-16 network model, optimize the fully connected layer from 3 to 2, and replace the original SoftMax classification layer with a 5-label SoftMax classifier.
- Fine-tune transfer learning. The VGG pre-training model parameters are used to optimize the parameters of the nutrient forage recognition model through migration learning, and the parameters of 13 convolutional layers and pooling layers are mainly determined.

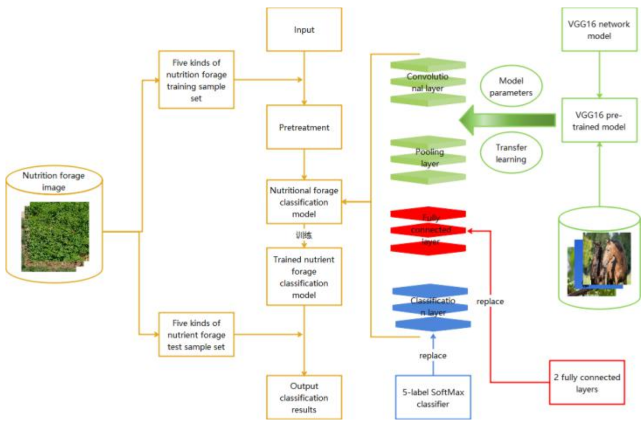


Fig.2 The framework of the nutrient forage classification model

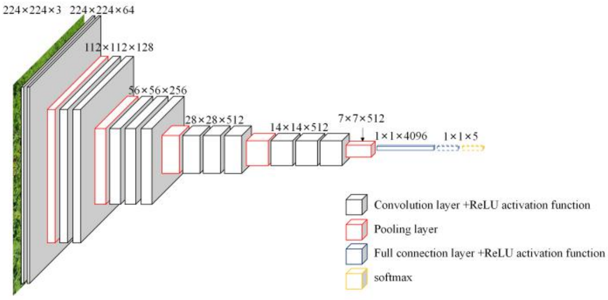
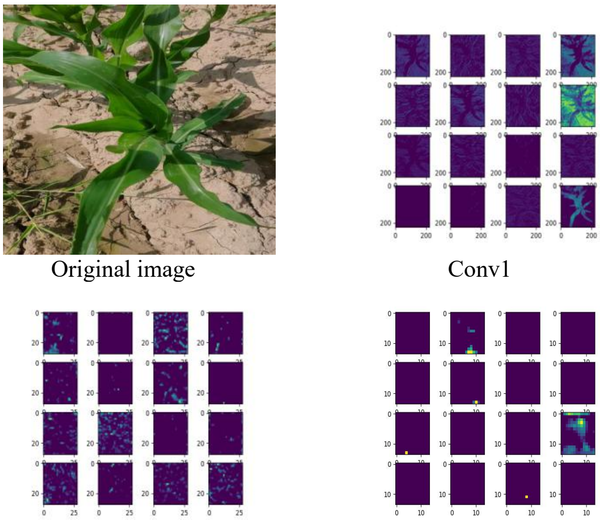


Fig.3 Improved VGG-16 network

4.EXPERIMENTTAL RESULTS AND ANALYSIS

4.1. Analysis of Feature Extraction Results of Model Convolutional Layer

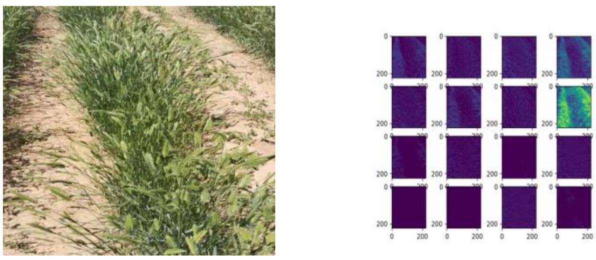
The feature maps of different convolutional layers of corn, wheatgrass, and alfalfa are shown in Figures 4, 5, and 6. The forage image is processed by Conv1 to get 64 feature maps, after Conv8 is processed, 512 feature maps are obtained, and after Conv13 is processed 512 feature maps, from the convolution feature image, it can be seen that Conv1 extracts the edge and texture features of the forage. Conv8 extracts more directional edge features than Conv1, highlighting the local features of the forage in multiple directions; Conv13 extracts the deeper level of the forage The characteristics of different pastures are more distinguishable. In contrast, the deeper the number of layers, the more representative the extracted features.



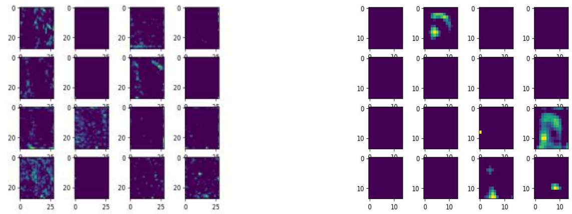
Original image Conv1

Conv8 Conv13

Fig. 4 Features of corn with different convolution layers

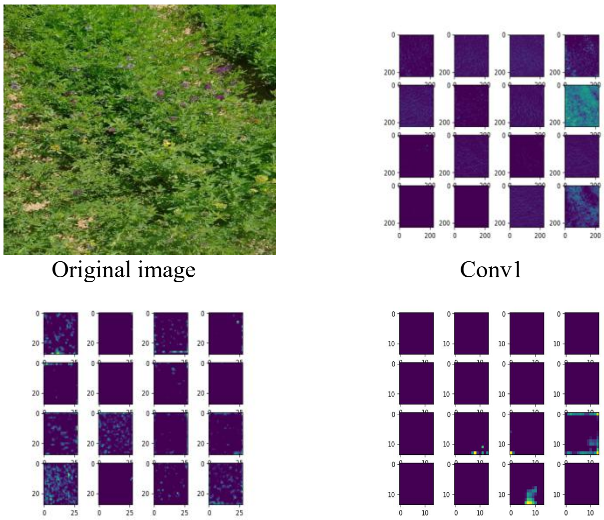


Original image Conv1



Conv8 Conv13

Fig.5 Feature map of different convolutional layers of Agropyron mongolia



Original image Conv1

Conv8 Conv13

Fig.6 Feature map of alfalfa with different convolutional layers

4.2. Model evaluation result analysis

Train 1753 training sets, 191 verification sets, and 100 test sets, which are divided into alfalfa, Mongolia wheatgrass, Nongke No. 1 tree Kochia, ribbed vetch, and corn. The model accuracy and loss function changes during the training process are shown in Figure 7. Choose Adam optimization algorithm, the convergence speed is fast, the learning rate is 0.001, training 50 epochs, the loss function has stabilized and reduced to the minimum, the training set accuracy can reach more than 95%, the training set loss is reduced to less than 0.2, in the training process There is no over-fitting phenomenon in.

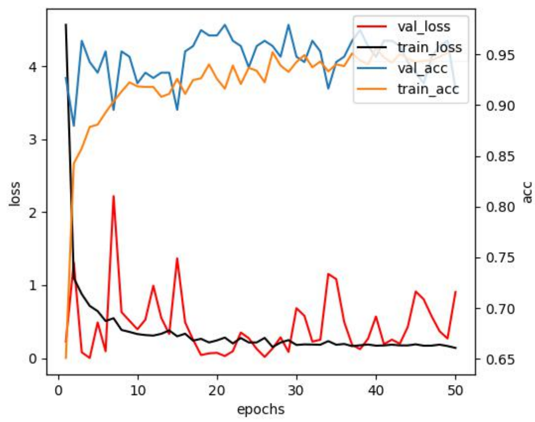


Fig.7 Accuracy rate and loss change curve of nutrient forage identification model

The confusion matrix [11] is used to measure the accuracy of the classifier's classification by comparing the predicted value with the actual value. The 100 test sets are tested in the training model, and various accuracy rates, recall rates and F-values are calculated respectively to objectively evaluate the accuracy of the model. The classification effect of the model is judged by the classification report and classification accuracy rate. Table 4 is the classification report of the test data. All types of F-values have reached more than 90%, balancing accuracy and recall rates, considering low-probability events, which can objectively express the pros and cons of the algorithm compared to the correct rate. The VGG-16 migration learning algorithm is used to analyze the nutritional forage image The classification achieves a higher classification accuracy.

Table 4 Classification report

	Accuracy	**Recall rate**	**Correct rate**	**F value**	**Test the total number of various images**
Alfalfa	90.91%	100%	90.91%	95.24%	20
Agropyron mongolia	100%	90%	90%	94.74%	20
Nongke No. 1 Tree Kochia	86.96%	100%	86.96%	93.03%	20
Rib Vicia	100%	85%	85%	91.89%	20
corn	100%	100%	100%	100%	20

Table 5 compares the classification accuracy of other algorithms with the VGG-16 algorithm. Use K-nearest neighbor algorithm (K-NN), support vector machine (SVM) and neural network algorithm (BP) to classify grass based on 15-dimensional features such as image texture and color, using the same data set as the VGG16 algorithm. The results show that the three algorithms have a low recognition rate in the recognition between alfalfa, Nongke No. 1 tree Kochia and ribbed pea. The VGG-16 migration learning algorithm used in this paper has high recognition accuracy for various types, and the recognition rate of each type is more than 90%.

Table.5 Comparison of accuracy rates of different classification algorithms

Classification algorithm	K-NN	SVM	BP	VGG-16
Correct rate(%)	60.1%	63.6%	69.8%	91.3%
Test time (s)	1.50	1.23	3.96	2.91

5. CONCLUSION

Based on the VGG-16 convolutional neural network, this research replaces the original 3 fully connected layers with 2 fully connected layers, and replaces the SoftMax classifier in the original VGG-16 network with a 5-label SoftMax classification layer. The classification and identification of 6 types of nutrient forages, including alfalfa, Mongolian wheatgrass, Nongke No.1 tree Kochia scoparia, ribbed pea, and corn, are classified and identified, and the following conclusions are obtained:

1) Pasture feature extraction experiments show that the weight parameters of the model convolutional layer and the pooling layer are shared through micro-transfer learning, which reduces network training parameters, and can effectively extract multi-layer feature images of pastures, and deepen the volume. The number of layers highlights the texture and edge information of the forage, and provides characteristic information for the classification and recognition of the forage. The forage classification method based on deep learning is more efficient than relying on manual direct observation.

2) The comparison experiment results of forage classification models show that the deep network model is better than the shallow network model. Compared with the K-NN, SVM, and BP neural network models, the recognition accuracy of the deep VGG-16 network is about 20% higher. Solve the problem of time-consuming, laborious and long cycle.

REFERENCES

[1] Qingke Wen,Zengxiang Zhang,Xiaoli Zhao,Ling Yi,Xiao Wang,Shunguang Hu,Bin Liu,Tian Zeng. Regularity and causes of grassland variations in China over the past 30?years using remote sensing data[J] . International Journal of Image and Data Fusion . 2015,6 (4) : 330-347.

[2] Li-Xia L，Gui-Min L，Yun-Xia L.Current Status and Prospect in the Biotechnology of Forage Breeding in China[J]. Acta Agrestia Sinica，2009，17(3) : 389-397.

[3] Wei Lifeng. Research on detection method of maize variety based on hyperspectral image[D].Shenyang Agricultural University，2017.

[4] Hu W，Huang Yangyu，Wei Li. Deep Convolutional Neural Networks for Hyperspectral Image Classification[J]. Journal of Sensors，2015，2015(2) : 1-12.

[5] Dey P, Banerjee N, Kaur R. Digital image classification with the help of artificial network by simple histogram[J]. Journal of Cytology，2016，33(2) : 63-65.

[6] Li Yandong, Hao Zongbo, Lei Hang. Review of Convolutional Neural Network Research[J]. Computer Applications, 2016, 36(9): 2508-2515.

[7] Simonyan K., Zisserman A.. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]. Computer Vision and Pattern Recognition, 2014.

[8] KETKAR N.Convolution neural networks.In deep with Python[M].Berkeley,CA:Apress,2017.

[9] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning [M]. Translated by Zhao Shenjian and others. Beijing: People's Posts and Telecommunications Press, 2017

[10] DENG J，DONG W，SOCHER R，et al.ImageNet : a Large-Scale Hierarchical Image Database[C]//2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR 2009). Miami，Florida，USA，2009 : 20-25.

[11] Liu Xiaoyi. Research and application of spatio-temporal model for sequence remote sensing image anomaly detection[D]. University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences), 2018.

[12] Wang Jingxuan, Feng Quan, Wang Yutong, Shao Xinqing. Research on the classification of legume pasture based on image recognition technology[J].Acta Grassland Sinica,2010,18(01):37-41.

[13] Han Ding, Wu Pei, Zhang Qiang, Han Guodong, Tong Fei. Feature extraction and image recognition of typical grassland pasture based on color moments[J].Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(23): 168-175.

[14] Han Ding,Wu Huijuan,Ma Ziyin,Korea Dong,Wu Pei,Zhang Qiang.Research on Feature Extraction and Recognition of Typical Pastures in Grassland[J].Chinese Journal of Grassland,2019,41(04):128-135.