Fine Classification of Typical Farms in Southern China Based on Airborne Hyperspectral Remote Sensing Images

1st Xin Hu
State Key Laboratory of Information Engineering in Surveying Mapping, and Remote Sensing
Wuhan University
Wuhan, China
whu_huxin@whu.edu.cn

2nd Yanfei Zhong
State Key Laboratory of Information Engineering in Surveying Mapping, and Remote Sensing
Wuhan University
Wuhan, China
zhongyanfei@whu.edu.cn

3rd Chang Luo
School of Geosciences an Info-physics
Central South University
Changsha, China
908474457@qq.com

4th Lifei Wei
School of Resources and Environmental Science
Hubei University
Wuhan, China
188233464@qq.com

***Abstract*—In the southern part of China, peculiar land fragmentation so that crop planting is characterized by small planting area of a single block, alternate cropping in multiple plots and diversified planting in space. Based on the unique crop planting characteristics in southern part of China, this paper take typical southern farm in Honghu City, Hubei Province as an example, adopting the platform of unmanned aerial vehicle (UAV) to carry hyperspectral imaging spectrometer to obtain the "double high" (high spectral and high spatial resolution) images at the same time. To complete the crop fine classification of 'double high' images , the CNN-CRF algorithm is proposed. The CNN-CRF algorithm acquires 91.5% accuracy with only 1% train samples on remote sensing images, which performs far better than most traditional classification approaches.**

Keywords*—*Airborne hyperspectral, Convolutional Neural Network (CNN), Conditional Random Fields (CRF), Fine Classification

INTRODUCTION

Obtaining and analyzing crop information quickly is the premise and foundation for precision agriculture practice[1], and crop acreage estimation is a core issue for analyzing crop information in farmland. In recent years, remote sensing technology has become the main way of crop acreage estimation because of its fast, simple, macroscopic, non-destructive and objective advantages. Aerospace and aerial remote sensing technology has been widely used in crop remote sensing classification. However, in the southern part of China, peculiar land fragmentation occurs due to historical reasons so that crop planting is characterized by small planting area of a single block, alternate cropping in multiple plots and diversified planting in space[2]. Based on the unique crop planting characteristics in southern China, remote sensing classification of crops first requires remote sensing images with high spatial resolution. Then, concerning that the spectra of different crops are very similar, remote sensing images are required to accurately define different crops with high spectral resolution. In short, the remote sensing image used for classification has high spatial resolution and spectral resolution simultaneously. However, remote sensing satellites cannot obtain the image with high spectral resolution and high spatial resolution simultaneously because of the limitation of imaging technology. And the way of hyperspectral imager carried by aircraft is very expensive.

Along with the development of the UAV technology currently, the remote sensing information acquisition technology of UAV is widely used in agriculture field because of its low operating cost, high flexibility and real-time data acquisition. UAV mounted hyperspectral imager can control flight height and acquire "double high" remote sensing images with high spectral resolution and high spatial resolution. However, there are several critical problems in the classification of hyperspectral data: 1) the serious phenomenon of 'same spectral from different materials' and 'same material with different spectral' among the crops appears; 2) curse of dimensionality, because of the high number of spectral channels[3]. Traditional classification model based on statistical pattern recognition is difficult to classify and recognize the hyperspectral data, and the most advanced classification algorithm is also needed.

In this paper, we take typical southern farm in Honghu City, Hubei Province as an example, adopting the platform of unmanned aerial vehicle (UAV) to carry hyperspectral imaging spectrometer to obtain the 'double high' (high spectral and high spatial resolution) images at the same time. The classification algorithm uses convolution neural network to automatically extract the high-level abstract features, and conditional random field is used for classification. The accuracy of classification algorithm is 91.79% under only 1% sample, which is far better than the traditional classification algorithm with remote sensing image.

I. METHODS

A. Convolutional Neural Networks（CNN）

CNN is a hierarchical framework utilizing the alternating convolutional layers, pooling layers, fully connected layers, et al. to automatically extract intrinsic features[4], which uses local connections to effectively extract the spatial information and shared weights to significantly reduce the number of parameters.

In order to avoid the size of the feature map from falling too fast , the pooling layer is removed in this paper , the feature is extracted from convolutional layer and fully

978-1-5386-5038-7 / 18 / $ 31.00©2018 IEEE

connected layer , and the softmax classifier is used to output the rule image .

The convolutional layer is the core building block of a Convolutional Network that does most of the computational heavy lifting, which computes convolution of the input feature maps with convolutional kernels. The i-th feature map of the l-th layer output is

$$y_i^l = \sum_j f(w_{i,j}^l * y_j^{l-1} + b_i^l) \quad (1)$$

Where y_j^{l-1} is the feature map of layer $l-1$ that connected to feature map y_i^l , $w_{i,j}^l$ is the convolutional kernel of y_j^{l-1} , b_i^l is bias of y_j^{l-1} . $f(\cdot)$ is nonlinear active function, the rectified linear units (ReLU) is used, $*$ denotes convolutional operator.

The fully connected layer can be seen as a special convolutional layer with the convolution kernel size is 1, and the network node is fully connected to the previous layer.

Softmax classifier is a generalization of logistic regression to the case where we want to handle multiple classes, which Calculation formula is

$$h(x) = \begin{bmatrix} p(y^i = 1 \mid x, \theta) \\ p(y^i = 2 \mid x, \theta) \\ . \\ . \\ p(y^i = k \mid x, \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x}} \begin{bmatrix} e^{\theta_1^T x} \\ e^{\theta_2^T x} \\ . \\ . \\ e^{\theta_k^T x} \end{bmatrix} \quad (2)$$

$h(x)$ denotes the probability vector of samples i belonging to several classes , and $p(y^i = k \mid x, \theta)$ is the probability of the i-th sample belonging to the k-th class.

B. Conditional Random Fields（CRF）

CRF is a probabilistic discriminative model based on statistic[5], which directly models the posterior probability of the labels given the image data $P(x \mid y)$ as follows.

$$P(x \mid y) = \frac{1}{Z(y)} \exp\left\{ -\sum_{c \in C} \psi_c(x_c, y) \right\} \quad (3)$$

Where x is the corresponding class labels for the entire image, y represents the observed data from the input image, $Z(y) = \sum_x \exp\left\{ -\sum_{c \in C} \psi_c(x_c, y) \right\}$ is the partition function, and $\psi_c(x_c, y)$ is the potential function.

The potential function based on the size of the variables in the cliques can be divided into unary potentials, pairwise potentials, and even high-order potentials, which corresponding energy functio in can be defined as follows:

$$E(x \mid y) = -\log P(x \mid y) - \log Z(y) = \sum_{c \in C} \psi_c(x_c, y) \quad (4)$$

CRF has been widely used in remote sensing image classification[6-8]. In the classification of image, the unary potential and pairwise potential are widely used. The unary potential and pairwise potential as shown in the Fig.1 and following formula:

$$E(x \mid y) = \sum_{i \in V} \psi_i(x_c, y) + \lambda \sum_{i \in V,\ j \in N_i} \psi_{ij}(x_i, x_j, y) \quad (5)$$

The unary potential ψ_i is a single point of potential energy which constructs the relationships between observation variables and the corresponding label field. The pairwise potential ψ_{ij} considers the spatial-contextual interaction between the observed variables and the label variable.

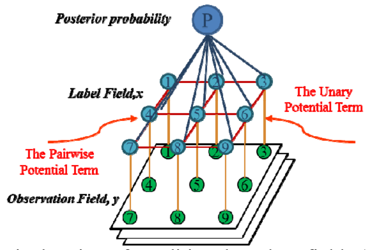


Fig.1 Schematic drawing of conditional random fields (CRF) in image classification.

According to the Bayesian maximum a posteriori (MAP) rule, the image classification aims to find the label image x by maximizing the posterior probability distribution function $P(x \mid y)$. Therefore, the MAP labeling x_{MAP} of CRF can be given by:

$$x_{MAP} = \arg\max_x P(x \mid y) = \arg\min E(x \mid y) \quad (6)$$

Therefore, maximizing the posterior probability distribution $P(x \mid y)$ is equivalent to minimizing the energy function $E(x \mid y)$.

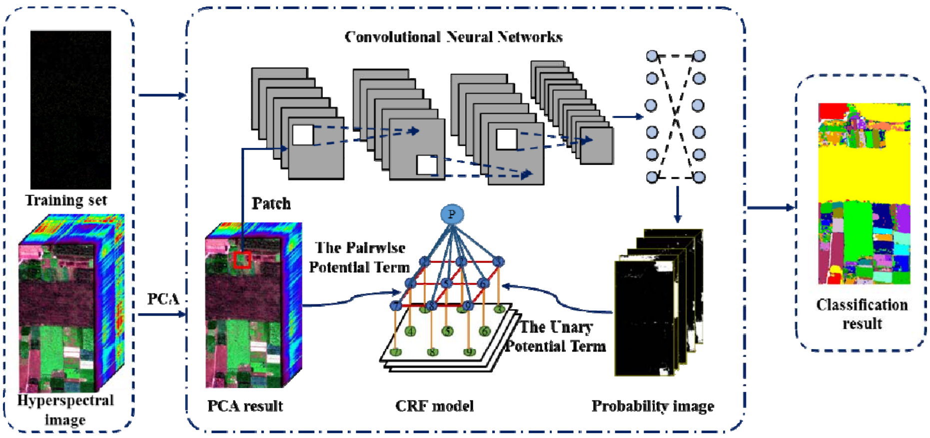


Fig.2 Flowchart of the crops fine classification based on Convolutional Neural Network (CNN) and Conditional Random Fields (CRF) algorithm

C. *Methodology of CNN-CRF*

In this paper, in order to make full use of the spatial-spectral information of hyperspectral remote sensing image, the CNN-CRF algorithm is proposed for fine classification of crops.The main procedure of the proposed method is illustrated in Fig. 2.

Firstly, PCA is introduced to condense the whole image, to reduce the data dimension to an acceptable scale and in the meantime reserving more than 99% of spectral information. Then, pick the patch as the training sample centered on the marked pixel, and the Convolutional Neural Networks is used to obtain rule image from the PCA image. Furthermore, the rule image as the input of unary potential, and the pairwise potential is modeled to consider the spatial interactions of the visible data in the neighborhood pixels, the CRF models will opuput the finall Classification map.

II. EXPERIMENTS AND ANALYSIS

A. *Data Description*

The WHU-Hi-Farmland-HH UAV dataset was acquired by a Headwall Nano-Hyperspec sensor mounted on an DJI M600 six-rotor unmanned autonomous vehicle (UAV) platform in November 2017. The experimental area is located in Honghu, Hubei province, China, which has an agricultural foundation with various crop species. The WHU-Hi-Farmland-HH UAV dataset can be used for precise agriculture mapping, and was provided by the Intelligent Data Extraction and Analysis of Remote Sensing group (RSIDEA group) of Wuhan University. The flight height of the UAV was 100 m, so that the spatial resolution of the acquired imagery is 4 cm. The image originally contained 940 × 475pixels and 274 spectral bands from the 400–1000 nm spectral range. In the data preprocessing, eight spectral channels were removed, leaving a total of 266 spectral bands to be used. An overview of this dataset and the distribution of the 22 corresponding semantic types are respectively given in Fig. 3(a) and (b). The 22 semantic types shown in Fig. 3(b) were labeled in detail based on field investigation, which covered almost the whole image. Fig. 3(c) presents some photos of the typical crop types obtained from the field investigation during the data collection. Table 1 reports the numbers of training and test samples in this image.

Table 1 Class information for the Hi-Farmland-HH UAV dataset

Class		**Samples**	
No.	Name	Train	Test
C1	Red roof	140	13901
C2	Road	35	3477
C3	Bare soil	218	21603
C4	Cotton	1633	161652
C5	Cotton firewood	62	6156
C6	Rape	446	44111
C7	Chinese cabbage	241	23862
C8	Pakchoi	41	4013
C9	Cabbage	108	10711
C10	Tuber mustard	124	12270
C11	Brassica parachinensis	110	10905
C12	Brassica chinensis	90	8864
C13	Small brassica chinensis	225	22282
C14	Lactuca saiva	74	7282
C15	Celtuce	10	992
C16	Film covered lettuce	73	7189
C17	Romaine lettuce	30	2980
C18	Carrot	32	3185
C19	White radish	87	8625
C20	Garlic sprout	35	3451
C21	Broad bean	13	1315
C22	Tree	40	4000

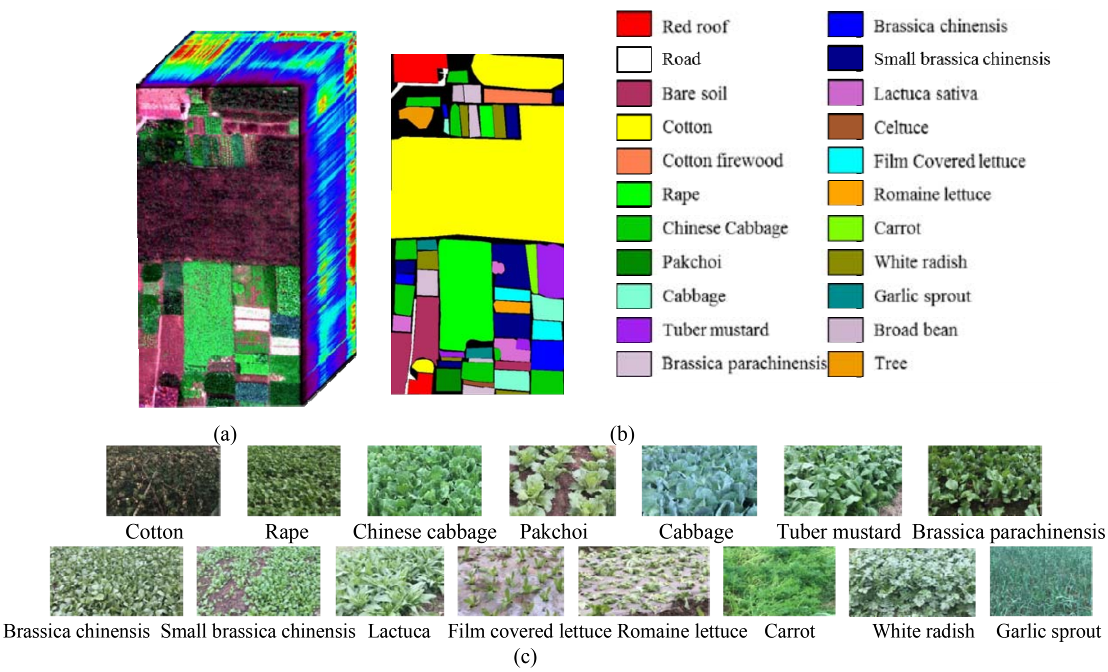
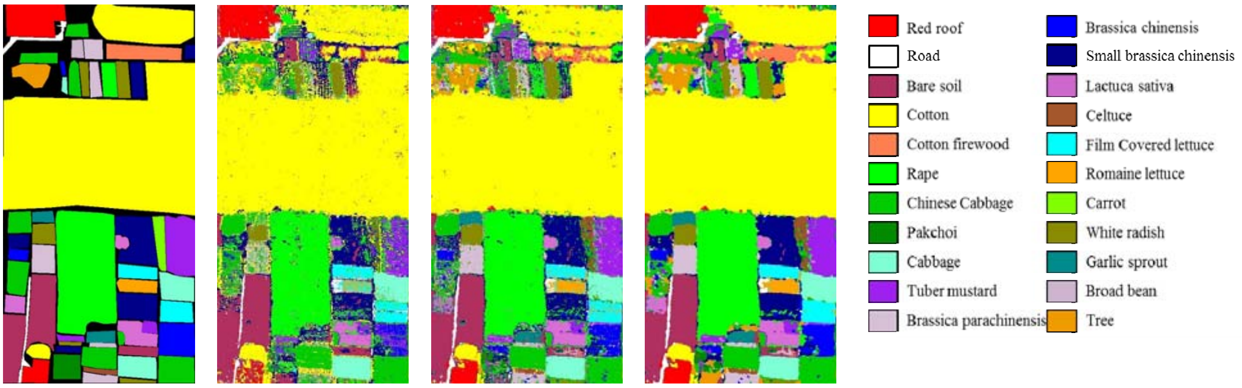


Fig. 3 Hi-Farmland-HH UAV dataset (a)Image cube. (b) Ground-truth image. (c) Typical crop photos in the study area

B. *Experiment Design and Results*

In order to retain 99% of the spectral information, we retained 20 principal components after PCA. Then, the patch size set as 5×5 to utilize the neighborhood information. The convolutional neural network consists of four convolution layers and two fully connected layers, and the size of the convolution kernel is 3×3. We use zero-padding operations in the first three convolution layers to prevent the



(a) Ground truth (b) SVM (c) CNN (d) CNN-CRF

Fig.4 The classification results for the Hi-Farmland-HH UAV dataset (a) Ground truth. (b)SVM. (c)CNN. (d) CNN-CRF.

size of feature map falling too fast. The parameters λ of CRF model set as 0.8.

To investigate the effectiveness of the proposed CNN-CRF algorithm, the CNN and SVM algorithm is used to compare experiments. The classification result are shown in Fig.4, and the corresponding quantitative performances are reported in Table 2.

Table 2 Classification accuracies for the Hi-Farmland-HH UAV dataset

Classifier	OA(%)	Kappa
SVM	80.43	0.7471
CNN	88.62	0.8557
CNN-CRF	91.79	0.8957

III. CONCLUSION

In this paper, targeting at typical fragmented and diversified planting pattern in southern China, we use airborne hyperspectral remote sensing images and use a combination of CNN and CRF classification algorithms to achieve much better classification accuracy than SVM, which can provide a reference for the fine classification of this diversified planting pattern.

ACKNOWLEDGMENT

I wish to thank Dr. Ji Zhao for guidance in CRF algorithm, and I wish to thank Ziran Yuan for assistance in data collection.

REFERENCES

[1] R. Earl, P. N. Wheeler, B. S. Blackmore, and R. J. Godwin, "Precision farming--the management of variability," vol. v. 51, 1996.

[2] W. Siming, "Study on the Fragmentariness of Land in China," *China Land Science,* 2008.

[3] G. Campsvalls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, pp. 1351-1362, 2005.

[4] Y. Lecun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard*, et al.*, "Handwritten Digit Recognition with a Back-Propagation Network," in *neural information processing systems*, 1990, pp. 396-404.

[5] J. D. Lafferty, A. Mccallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," in *international conference on machine learning*, 2001, pp. 282-289.

[6] J. Zhao, Y. Zhong, H. Shu, and L. Zhang, "High-Resolution Image Classification Integrating Spectral-Spatial-Location Cues by Conditional Random Fields," *IEEE Transactions on Image Processing,* vol. 25, pp. 4033-4045, 2016.

[7] J. Zhao, Y. Zhong, T. Jia, X. Wang, Y. Xu, H. Shu*, et al.*, "Spectral-spatial classification of hyperspectral imagery with cooperative game," *Isprs Journal of Photogrammetry and Remote Sensing,* vol. 135, pp. 31-42, 2018.

[8] Y. Zhong, T. Jia, J. Zhao, X. Wang, and S. Jin, "Spatial-Spectral-Emissivity Land-Cover Classification Fusing Visible and Thermal Infrared Hyperspectral Imagery," *Remote Sensing,* vol. 9, p. 910, 201