URBAN LAND USE/LAND COVER CLASSIFICATION BASED ON FEATURE FUSION FUSING HYPERSPECTRAL IMAGE AND LIDAR DATA

Qiong Cao, Yanfei Zhong, Ailong Ma, Liangpei Zhang*

The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China;
*Corresponding author E-mail: zhongyanfei@whu.edu.cn Phone +86-27-68779969

ABSTRACT

Hyperspectral images have been widely used in classification because of the abundant spectral information. But it can't distinguish the objective with similar spectral character but different elevation. However, LiDAR data can obtain elevation information. Therefore, it will obtain better classification maps if fusing the two data. In recent years, CNN has attracted much attention due to its powerful ability to excavate the potential representation and features of the raw data. However, it's difficult to distinguish the objects with different spectral information but similar surface character. Unlike CNN features, the traditional manual features, such as the normalized vegetation index (NDVI), have a certain characteristic expression significance. In order to consider both the semantic information of traditional manual features and the advanced features of CNN features, this paper proposes a fusion algorithm of hyperspectral and LiDAR fusion based on feature fusion. The proposed algorithm has achieved a good fusion classification effect on the MUUFL Gulfport Hyperspectral and LiDAR Data set.

Index Terms**—** Hyperspectral, LIDAR, land-use/land-cover classification, feature fusion, deep learning, convolutional neural network

1. INTRODUCTION

As an important technology, remote sensing has been widely used in the study of urban land use/land cover (LU/LC) classification. And with the continuous development of earth observation technology, there are a variety of remote sensing sensors [1] which can obtain different kinds of information. For example, high resolution image that can obtain abundant spatial information, multispectral and hyperspectral sensors that can get spectral characteristics, LiDAR sensors that can get accurate 3D information and SAR sensors that can obtain amplitude and phase information. The images obtained by these sensors have different advantages in the classification and other applications of urban area. However, there are shortages in the using of single source images. For example， hyperspectral images can characterize the spectral and spatial characteristics of objects [2], which is suitable for urban LU/LC classification. However, it's difficult to distinguish the objects with similar spectral characteristics but different elevation information, such as roof and road which are both made of concrete. Compared with hyperspectral images, LiDAR data has accurate three-dimensional information [3], it can classify the objects by height. However, since it lacks the semantic information of objects, it has poor ability to classify the objects with similar elevation and different spectral information. For example, two roads with the same height but made of different materials. Therefore, it will greatly improve classification result by fusing the two types of data [4].

Data fusion is divided into three levels, pixel level, feature level and decision level. Feature level fusion refers to the combination of features extracted from different images to form a new set of features. It has been widely used in the fusion of hyperspectral and LiDAR data because of its small amount of computation and the effectiveness of feature extraction. In feature extraction, there are several main methods: (1) Methods based on segmentation [5]. The results of the segmentation form as the result of feature extraction. (2) Methods based on morphological operation [6]-[7]. Extracting the features by morphological operation. (3) Methods based on deep learning. Using deep learning model as a feature extractor to obtain the features [8]-[9].

In recent years, deep learning has attracted much attention due to its powerful ability to excavate the potential representation and features of the raw data [10]. The convolution neural network (CNN) simulates the concept of "local visual field" in the human visual system, and transforms the full connection into a local connection for reducing the computational cost and considering the local spatial information. In addition, it is widely used in image classification because of its ability to learn rich hierarchical representations and the automation of the learning process. Because CNN can extract the suitable features according to different source data, it is appropriate for the classification fusing hyperspectral image and LiDAR data. However, the high-level features extracted by CNN often do not have any representative meaning and lack semantic information. For hyperspectral images, the CNN features lack of spectral information which is most meaningful to the classification. Unlike CNN features, the traditional manual features, such as the normalized vegetation index (NDVI), the principal component analysis (PCA) features, have a certain characteristic expression significance. Therefore, this paper proposed the combination of traditional manual features and CNN features to improve the classification effect.

In this paper, a feature fusion framework (Fig. 1) is proposed to deal with the elaborate classification of an urban area using hyperspectral image and LIDAR data. In the proposed framework, manual features were first extracted from hyperspectral image and LiDAR data respectively, get the spectral and spatial characteristics of hyperspectral image, and the elevation and intensity characteristics of LiDAR data. Then extracting the CNN features from the hyperspectral raw data and the LiDAR features (intensity map and the DSM map) by CNN model. Finally, fusing the manual features and CNN features to make full use of the complementarity between the two features. And the final classification map was obtained by classifying the fused features. The proposed algorithm has achieved a good fusion classification effect on the MUUFL Gulfport Hyperspectral and LiDAR Data set [11]-[12].

2. THE FEATURE FUSION FRAMEWORK

The main procedure of the data fusion framework (Fig. 1) is described as follows:

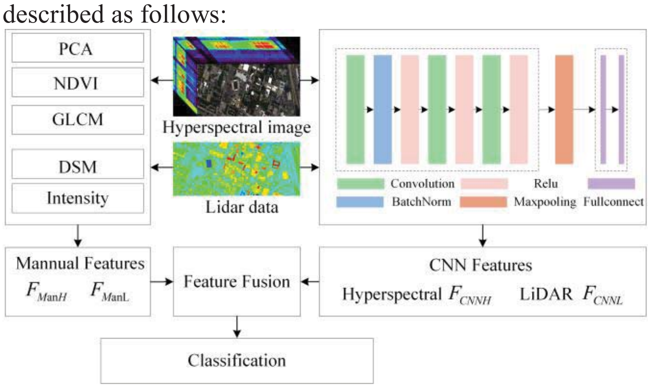


Fig. 1. Proposed framework for the feature fusion of hyperspectral imagery and LIDAR data.

2.1 Manual Feature Extraction

In order to improve the separability of the class, the spectral and spatial features of hyperspectral images and the elevation and intensity characteristics of LiDAR data were extracted.

(1) Spectral features extraction. In order to reduce the complexity of the computation, the feature reduction algorithm was used to reduce the spectral features. In this paper, PCA was used to obtain the feature F_{PCA} . In order to improve the separability of vegetation, NDVI was used to extract the feature F_{NDVI} .

(2) Spatial features extraction. In the image classification task, the spatial information of remote sensing images is very important. The gray level co-occurrence matrix is obtained by computing the image, and then some spatial texture features of the image can be calculated by computing the partial eigenvalues of the co-occurrence matrix. Homogeneity can measure the local changes in the texture of the image. The large value indicates that there is no change between the different regions of the image texture. Therefore, in order to increase the separability of the class, the spatial texture feature F_{GLCM} was extracted by the homogeneity characteristics of the grayscale symbiotic matrix from hyperspectral image.

(3) LiDAR feature extraction. In order to measure the difference in the height of different objects, digital surface model (DSM) was used to represent the elevation feature F_{DSM} . The intensity information of different ground objects is different when they reflect the echo signal. Therefore, the intensity map (i.e., F_{Inten}) of LiDAR data was used to represent the differences of materials.

The final manual features of hyperspectral image and LiDAR data was F_{ManH} and F_{ManL} .

$$F_{ManH} = F_{PCA} + F_{NDVI} + F_{GLCM} \quad (1)$$

$$F_{ManL} = F_{DSM} + F_{Inten} \quad (2)$$

2.2 CNN Feature Extraction

CNN are a type of feed-forward artificial neural network containing at least one convolutional layer in nets that conducted convolution computation [13]. Compared with traditional manual features, CNN provides a simple and automatic training method. In this paper, CNN was used to extract hierarchical features from raw hyperspectral image and LiDAR manual features respectively.

(1) Pre-processing. The image I consists of K×K neighborhoods of a pixel in the raw data was used to as the input of the CNN model. And in this paper, the input image I will be normalized to [0,1] to accelerate the learning convergence by being divided by the maximum pixel value:

$$I' = \frac{I}{Max} \quad (3)$$

Where *I'* is normalized image and Max is the maximum pixel value in the image *I*.

(2) Feature training. After pre-processing, the CNN model was used to extract the hierarchical features, which consists of convolutional layers, batch norm layer, Relu layers, pooling layer, fully connectedly layer and SoftMax layer i.e. classification layer. Batch norm layer and Relu layers were used to accelerate the learning convergence and prevent overfitting. Pooling layers provides a strong robustness and reduces the number of parameters to prevent the overfitting. In this paper, max-pooling was used, which keeps the maximum value in a small area. The fully connected layer connects the classification layer to map the distributed feature representation to the sample space.

(3) Feature extraction. The model was trained by multiple feedback adjustment using training samples, then the prediction results was obtained by predicting the whole image. In this paper, two fully connectedly layer i.e. *fc1* and *fc2* were used, the *fc1* layer in the prediction network was extracted as a CNN feature. The CNN features F_{CNNH} and F_{CNNL} were obtained by training raw hyperspectral image and LiDAR manual features respectively.

2.3 Feature Fusion and Classification

In order to make full use of the complementarity of the manual features and CNN features, these two features were fused to obtain a new feature set. According to the equation (3), the extracted two features were normalized respectively, and then stacking these features to get the fused characteristics F_{Fused} .

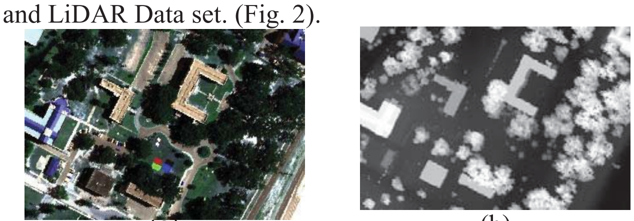
$$F_{Fused} = F_{ManH} + F_{CNNH} + F_{ManL} + F_{CNNL} \tag{4}$$

The support vector machine (SVM) [14] was used to classify the fused features. SVM is established based on the Vapnik-Chervonenkis (VC) dimension theory and risk minimization principle, which has many unique advantages in solving small-sample, nonlinear and high-dimensional pattern recognition. Therefore, the SVM was chosen to obtain the final classification map.

3. EXPERIMENT AND ANALYSIS

3.1 Data sets

The experiment utilized the MUUFL Gulfport Hyperspectral and LiDAR Data set. (Fig. 2).



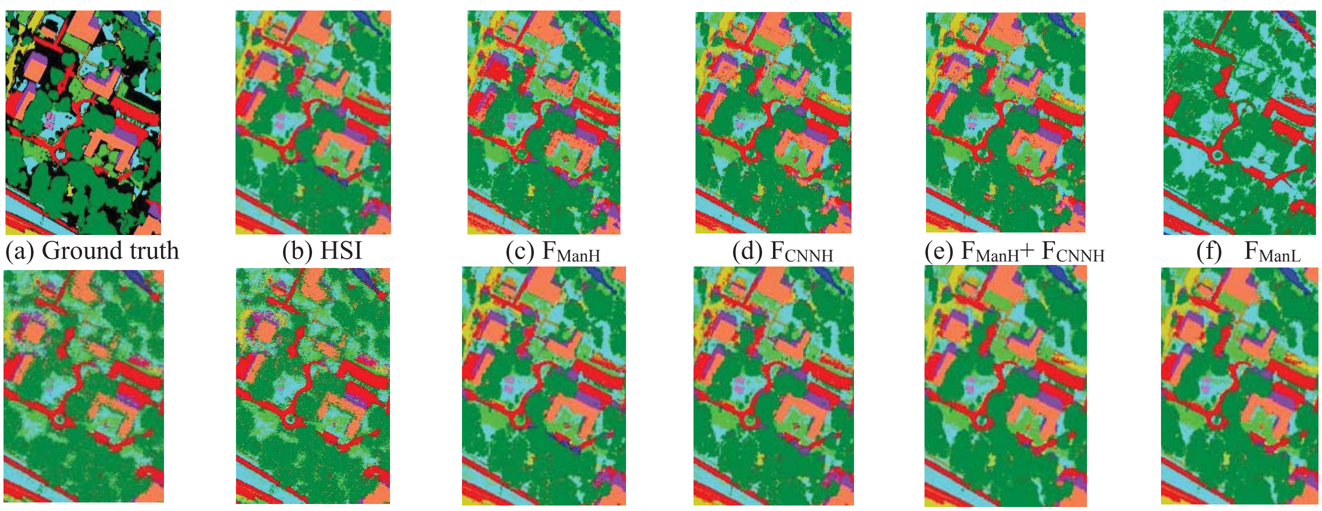
(a) (b)

Fig. 2. Experiment data. (a) hyperspectral image. (b) DSM.

The hyperspectral imagery has 64 spectral bands with 10 nm spectral resolution. The size of hyperspectral image and DSM data are both 325×220 with 1 m spatial resolution. The MUUFL Gulfport data set was collected in November 2010 over the University of Southern Mississippi Gulf Park Campus, located in Long Beach, Mississippi. Each pixel in the image will be mapped to one of 11 classes.

3.2 Experiment design

The parameters involved in the experiment are as follows: the number of reserved bands after PCA is 6, GLCM is generated by the first three principal components of PCA with the window size 9×9, the input image size of the CNN model is 9×9, the learning rate is set to 0.001, and the extracted CNN feature dimension is 64.



(a) Ground truth (b) HSI (c) F_{ManH} (d) F_{CNNH} (e) F_{ManH}+ F_{CNNH} (f) F_{ManL}

(g) F_{CNNL} (h)F_{ManL}+F_{CNNL} (i) HSI+ F_{ManL} (j) F_{ManH}+F_{ManL} (k)F_{CNNH}+F_{CNNL} (l) F_{Fused}

Fig.3. Final classification map

3.3 Experiment result and analysis

As shown in the final classification map, i.e. fig.3, the result of feature fusion is less noisy and has the best visual effect. For fig. (f) and (g), it is obvious that the classification was improved by adding CNN features for LiDAR data. The final overall classification accuracy is 82%. The classification accuracy of each category is shown in Table 1.It can be seen from both the precision of each class and the overall accuracy that the proposed algorithm has achieved a better classification effect. (1) Comparing fifth columns, eighth columns and twelfth columns in Table 1, we can see that the fusion of hyperspectral image and LiDAR data has significantly improved the classification accuracy, and there is no decline of any category. Especially in buildings and roads, the effect of classification has been obviously improved after adding LiDAR data. (2) For classification results using CNN features and manual features, we can see that for most classes CNN features are superior to manual features, except for some categories that require higher spectral information, such as Yellow curb and Cloth panels. Meanwhile, it also shows that the fusion of manual features and CNN features is necessary. (3) For the classification results using CNN features and fused features, we can see that the accuracy of all categories and the overall accuracy have been improved, especially for the mentioned yellow curb and cloth panel which classification effect has been significantly improved.

Table 1. Classification accuracy

	HSI	F_{ManH}	F_{CNNH}	F_{ManH}+ F_{CNNH}	F_{ManL}	F_{CNNL}	F_{ManL}+ F_{CNNL}	HSI+ F_{ManL}	F_{ManH}+ F_{ManL}	F_{CNNH}+ F_{CNNL}	F_{Fused}
Trees	94.45	95.71	96.00	96.70	84.93	88.74	88.84	96.27	96.85	97.48	97.65
Mostly grass	64.87	67.19	70.91	74.97	0	37.52	38.15	71.09	72.31	74.55	76.87
Mixed Surface	81.05	81.45	78.93	81.06	73.15	65.22	64.93	81.34	82.21	81.58	82.63
Dirt/sand	81.38	81.44	79.00	79.49	0	29.34	30.49	81.01	81.01	85.21	85.70
Road	80.49	80.69	88.77	89.18	83.75	86.44	87.11	92.46	92.37	93.90	94.17
Water	77.33	73.75	37.71	63.48	53.46	60.14	60.62	83.29	79.47	89.98	90.93
Building shadow	67.16	66.62	77.06	78.56	0	29.80	30.65	66.42	69.55	80.60	80.75
Buildings	79.59	76.07	90.19	90.58	0.61	76.75	77.19	93.29	94.71	96.08	96.49
Sidewalk	38.04	49.68	50.24	54.33	0	27.13	27.21	43.58	55.70	61.80	62.84
Yellow curb	50.30	48.48	6.67	15.15	0	12.73	12.73	51.52	46.06	35.15	36.97
Cloth panels	87.60	89.67	49.17	61.57	0	28.10	28.10	88.43	88.84	71.49	73.55
OA(%)	75.19	75.77	77.97	79.36	51.41	65.79	66.03	79.29	80.20	81.50	**82.00**

4. CONCLUSION

In order to take account of both the semantic information of manual features and the high-level information of CNN features, a hyperspectral and LiDAR fused classification algorithm based on feature fusion is proposed in this paper. The algorithm has achieved good result on the MUUFL Gulfport data set. The fusion of two kinds of data and the fusion of two set of features both greatly improve the classification effect.

REFERENCES

[1] R. Momeni, P. Aplin, D. Boyd, "Mapping Complex Urban Land Cover from Spaceborne Imagery: The Influence of Spatial Resolution, Spectral Band Set and Classification Approach," Remote Sens., vol. 8, no. 2, pp. 88, 2016.

[2] JA. Benediktsson, JA. Palmason, JR. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," IEEE Trans. Geosci. Remote Sens., vol. 43, no. 3, pp. 480-491, 2005.

[3] J. Zhang, X. Lin, X. Ning, "SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas," Remote Sens., vol. 5, no. 8, pp. 3749-3775,2013, 5(8).

[4] Y. Zhong, Q. Cao, J. Zhao, A. Ma, B. Zhao, L. Zhang, " Optimal Decision Fusion for Urban Land-Use/Land-Cover Classification Based on Adaptive Differential Evolution Using Hyperspectral and LiDAR Data," Remote Sens., vol. 9, no. 8, pp. 868,2017.

[5] A. Merentitis, C. Debes, R. Heremans, N. Frangiadakis, "Automatic fusion and classification of hyperspectral and LiDAR data using random forests," in Proc. IEEE Geoscience and Remote Sensing Symposium, pp. 1245-1248,2014.

[6] Ghamisi P, Benediktsson J A, Phinn S, "Fusion of hyperspectral and LiDAR data in classification of urban areas," in Proc. IEEE Geoscience and Remote Sensing Symposium, pp. 181-184,2014.

[7] Liao W, Pižurica A, Bellens R, et al, "Generalized Graph-Based Fusion of Hyperspectral and LiDAR Data Using Morphological Features," IEEE Geosci. Remote Sens. Lett., vol. 12, no. 3, pp. 552-556,2015.

[8] Chen Y, Li C, Ghamisi P, Jia X, Gu Y, "Deep Fusion of Remote Sensing Data for Accurate Classification," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 8, pp. 1253 - 1257, 2017.

[9] Chen Y, Li C, Ghamisi P, et al, "Deep fusion of hyperspectral and LiDAR data for thematic classification," in Proc. IEEE Geoscience and Remote Sensing Symposium, pp. 3591-3594, 2016.

[10] Xu X, Li W, Ran Q, et al, "Multisource Remote Sensing Data Classification Based on Convolutional Neural Network," IEEE Trans. Geosci. Remote Sens., vol. pp, no. 99, pp. 1-13.

[11] P. Gader, A. Zare, R. Close, J. Aitken, G. Tuell, "MUUFL Gulfport Hyperspectral and LiDAR Airborne Data Set," University of Florida, Gainesville, FL, Tech. Rep. REP-2013-570, Oct. 2013.

[12] X. Du and A. Zare, "Technical Report: Scene Label Ground Truth Map for MUUFL Gulfport Data Set," University of Florida, Gainesville, FL, Tech. Rep. 20170417, Apr. 2017. Available: http://ufdc.ufl.edu/IR00009711/00001

[13] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Book in preparation for MIT Press.

[14] Cortes, C, Vapnik, V, "Support-Vector Networks," Mach. Learn. vol. 20, pp. 273–297,1995.