COLLABORATIVE CLASSIFICATION OF HYPERSPECTRAL AND LIDAR DATA WITH INFORMATION FUSION AND DEEP NETS

Chen Chen [1], *Xudong Zhao* [2], *Wei Li* [2], *Ran Tao* [2], *Qian Du* [3]

[1] College of Information Science & Technology, Beijing University of Chemical Technology
[2] School of Information and Electronics, Beijing Institute of Technology, Beijing, China
[3] Department of Electrical and Computer Engineering, Mississippi State University

ABSTRACT

Convolutional neural network (CNN) receives extensive attention in hyperspectral image classification. While hyperspectral images contain abundant spectral information but lack spatial information, which usually contributes to poor classification results. In this paper, a novel classification framework called information fusion based CNN (IF-CNN) is proposed to compensate for the shortcomings of hyperspectral images. The proposed method merges hyperspectral images with abundant spectral information and LiDAR images with rich spatial information as the input of classification framework. Furthermore, the framework consists of two convolutional neural networks: one-dimensional CNN for extracting spectral features, and two-dimensional CNN for extracting spatial correlation features. Experimental results demonstrate that the proposed method achieves excellent performance compared with some existing methods.

***Index Terms*—** Hyperspectral Image, Information Fusion, Convolutional Neural Network, Deep Learning, Pattern Recognition.

1. INTRODUCTION

Hyperspectral imagery (HSI) attaches considerable inerest in recent years and has been widely applied in mineral exploitation, environmental science and earth observation, etc. In the early stage of HSI classification, many machine learning algorithms were introduced. Support vector machine (SVM) [1], extreme learning machine (ELM) [2], relevance vector machine (RVM)[3] and markov random fileds (MRFs)[4] were investigated and achieved satisfactory performance.

Recently, deep learning methods have been successfully applied in remote sensing image analysis. Zhang et al.[5] first attempted to evaluate the performance of many state-of-the-art deep learning algorithms on remote sensing images. Hu et al. [6] employed convolutional neural network to extract the spectral features of HSI, and the performance was obviously superior to SVM. Subsequently, further improvement researches based on CNN emerged. Diverse region-based CNN (DR-CNN) encode semantic context-aware representation to obtain promising features [7]. Pixel-Pair Features CNN (PPF-CNN) significantly increased the number of training samples by pixel pairing [8]. Two-Branch CNN adopted feature fusion method to fuse HSI and LiDAR or high-resolution visual image, which effectively combined the advantages of images from different sensors [9].

Hyperspectral images consist of hundreds of narrow spectral bands, in which rich spectral information helps to distinguish different materials. However, the low resolution of HSI limits its development in classification tasks. In this paper, we merge high-resolution LiDAR and HSI using information fusion method, and the merged images possess both high-resolution and rich spectral information. Furthermore, dual-tunnel CNN framework is used for classification. One-dimensional CNN and two-dimensional CNN are employed to extract spectral and spatial features, respectively.

2. INFORMATION FUSION BASED CNN FRAMEWORK

The flowchart of proposed IF-CNN classification framework is shown in Fig. 1. Part I. explains the information fusion process, Part II. illustrates dual-tunnel CNN classification framework, which consists of one-dimensional CNN and two-dimensional CNN.

HSI has large number of channels, which is quite useful for distinguishing different materials. However, the low resolution of HSI limits its development in classification tasks. While a prime feature of LiDAR is the high resolution, which is dominant in identifying elevation information. Therefore ,we use information fusion method to merge HSI and LiDAR, and obtain the image combining the advantages of both. The methodology we used is the Principal Component (PC) Spectral Sharpening in ENVI. ENVI is deployed a variety of information fusion methods. Specifically, PC Spectral Sharpening is especially suitable for processing HSI.

This work was supported by the National Natural Science Foundation of China under Grants NSFC-91638201, 61571033 (Corresponding author: Wei Li; Email: liwei089@ieee.org)

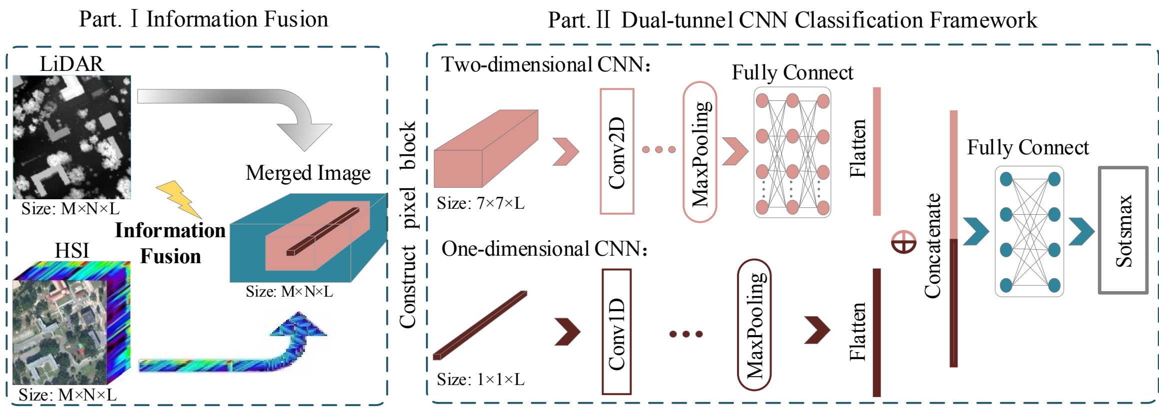

Fig. 1. Flowchart of the proposed IF-CNN for HSI classication.

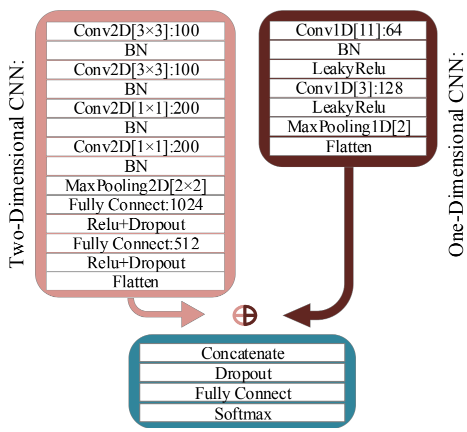

Fig. 2. The detailed structure of dual-tunnel CNN Classification Framework.

The concrete process of PC Spectral Sharpening method is as follows. 1) Performing principal component analysis on HSI and find each principal component. 2) Making a histogram to match LiDAR with the first principal component. 3) Replacing the first principal component with a high spatial resolution image produced in the previous step. 4) The new first principal component is inversely transformed with other principal components to generate merged images with high spatial resolution.

Dual-tunnel CNN classification framework is then used to extract features of the merged image. Both spectral and spatial information are critical to HSI pixel level classification. Therefore, we construct $7 \times 7 \times L$ pixel blocks for two-dimensional CNN, and $1 \times 1 \times L$ pixel blocks for one-dimensional CNN (L is the number of bands). Fig.2 demonstrates detailed structure of the framework. In two-dimensional CNN, each convolution process involves a certain operation, including two dimensional CNNs (conv2D) and two batch normalization layers. All the convolution operations are executed with zero padding, and the convolution stride is set as 1. In addition, Relu is employed as an activation function, Dropout is used to prevent overfitting. In one-dimensional CNN, one dimensional CNN (conv1D) is applied to extract spectral features, and convolution operations are executed without zero padding. Finally, the features obtained by the two CNNs are flattened and concatenated together. We use the fully connected network to extract features again and apply softmax to predict the classification label of the testing pixels. The learning rate is one of the factors that affects the convergence of speed and training performance. The learning rate is set as 0.0001 with the policy of Adam.

Table 1. Overall classification accuracy (%) versus different types of data. (H: HSI, L: LiDAR)

	MUUFL Gulfport				
	H	L1	L2	H+L1	H+L2
OA	88.90	67.91	70.00	90.64	**92.51**

3. EXPERIMENTAL RESULTS

For the proposed IF-CNN, all the programs are implemented in Python language, and the network is constructed using TensorFlow[1] and Keras[2]. TensorFlow is an open source software library for numerical computation using data flow graphs, and Keras can be seen as a simplied interface to TensorFlow.

3.1. Experimental Data

Classification performance of the proposed IF-CNN is evaluated on MUUFL Gulfport data. MUUFL Gulfport data

[1]http://tensorow.org/
[2]https://github.com/fchollet/keras

Table 2. Comparison of the overall classification accuracy (%) among the proposed method and the baselines using the MUUFL Gulfport data. (H: HSI, L: LiDAR)

class (Train/Test)	SVM (H)	SVM (H+L2)	ELM (H)	ELM (H+L2)	PPF-CNN (H)	PPF-CNN (H+L2)	Two-branch CNN (H)	Two-branch CNN (H+L2)	IF-CNN (H+L1)	IF-CNN (H+L2)
Trees (150/23096)	82.50	83.16	79.26	83.61	86.92	91.76	91.75	92.03	92.47	95.46
Mostly grass (150/4120)	78.57	77.16	77.55	82.33	81.84	75.85	86.31	75.32	92.84	85.44
Mixed ground surface (150/6732)	72.04	69.24	67.86	68.64	86.96	88.67	80.94	84.02	78.19	86.22
Dirt/Sand (150/1676)	86.04	86.16	83.47	86.93	93.26	88.90	92.84	95.41	98.45	97.02
Road (150/6537)	86.35	87.82	87.12	87.73	90.09	92.12	87.49	92.73	87.85	90.55
Water (150/316)	94.30	99.68	92.09	97.15	99.37	100.00	99.68	100.00	100.00	99.68
Building shadow (150/2083)	88.00	92.70	84.16	89.29	91.07	91.07	92.80	91.98	93.81	92.75
Buildings (150/6090)	78.19	89.43	80.57	89.11	91.49	94.24	87.70	94.01	96.45	95.14
Sidewalk (150/1235)	73.68	74.98	72.63	74.01	71.74	71.98	85.26	76.28	83.56	83.97
Yellow curb (150/33)	100.00	100.00	93.94	96.97	90.91	81.82	96.97	100.00	96.97	96.97
Cloth panels (150/119)	98.32	99.16	98.32	98.32	97.48	97.48	98.32	96.64	98.23	96.64
OA	81.06	82.63	79.09	82.97	87.57	89.90	88.90	89.79	90.64	**92.51**

set was collected in November 2010 over the University of Southern Mississippi Gulf Park Campus, located in Long Beach, Mississippi. The data collection contains coregistered hyperspectral and LiDAR data over the campus. The hyperspectral imagery consists of 325×220 pixels and 64 spectral channels. In additon, there are 11 different landcover classes in the ground truth.

3.2. Classification Performance

MUUFL Gulfport data contains one hyperspectral data and two LiDAR datas from two different flights. We mark hyperspectral image as H, LiDAR from two different flights as L1 and L2, the image generated by information fusion of H and L1 as H+L1, and the image generated by information fusion of H and L2 as H+L2. The classification results of the proposed method using these images respectively are illustrated in Table 1. The classification accuracies of H+L1 and H+L2 are significantly better than H, L1 and L2, which strongly suggests that the information fusion process combines the advantages of HSI and LiDAR. In addition, the result of L2 is superior to L1, and the result of H+L2 is superior to H+L1. Therefore, L2 is finally selected for information fusion with H, and H+L2 is utilized in the following experiments.

When the block window size is set to 11×11, the performance of the proposed IF-CNN is compared with some state-of-the-art HSI classication approaches, such as SVM, ELM, PPF-CNN and Two-branch CNN. In each class, we randomly select 150 samples for training and the rest for testing. The classification results are listed in Table 2. The classification with information fusion are obviously better than those with only HSI, and the proposed IF-CNN is significantly superior to all the other classifiers. Among them, the classification accuracies of Trees, Buildings and Road are significantly improved in each classifiers after information fusion. This is because that HSI mainly distinguishes different classes by spectral information of different materials. However, for some classes with similar materials, HSI is not advantageous in classification, such us trees and mostly grass, buildings and road. The elevation information of LiDAR can be used to solve this problem. There are many gaps between leaves of trees, and LiDAR can capture this feature to distinguish it from other classes; Buildings and road are all made of concrete, but the elevation of buildings is higher than road, which make it easy to distinguish them. Fig. 3 illustrates the classication maps, in which we present several classication results in Table 2. It can be easily find that the classication map achieved by H+L is evidently less noisy than that achieved by H. And IF-CNN(H+L2) achieves the best classification results.

The construction of pixel block facilitates pixel classification, since the spatial information contained in the surrounding pixels is critical to the central pixel. When the number of training samples is 150 per class, we compare the performance of two-branch CNN and IF-CNN on different pixel window sizes. The performances are demonstrated in Fig. 4 (a). Obviously the proposed IF-CNN performs better and the window size of 11×11 provides a more prominent performance. As shown in Fig. 4 (b), when the pixel window size is 11×11, we compare the classification of different numbers of training samples per class. As the number of training samples increases, the classification results are getting better. IF-CNN is superior to two-branch CNN and the maximum accuracy is achieved when the number of training samples is 150.

4. CONCLUSIONS

In this paper, a novel information fusion based CNN framework is proposed for hyperspectral image classication on MUUFL Gulfport data. The main contributions are summarized as follows. (1) IF-CNN is employed to combine the advantages of HSI and LiDAR images. LiDAR images have high resolution, strong penetration and sufficient elevation informations, which significantly aids HSI to achieve better classification results. (2) A dual-tunnel CNN classification framework is utilized to extract the features of the image

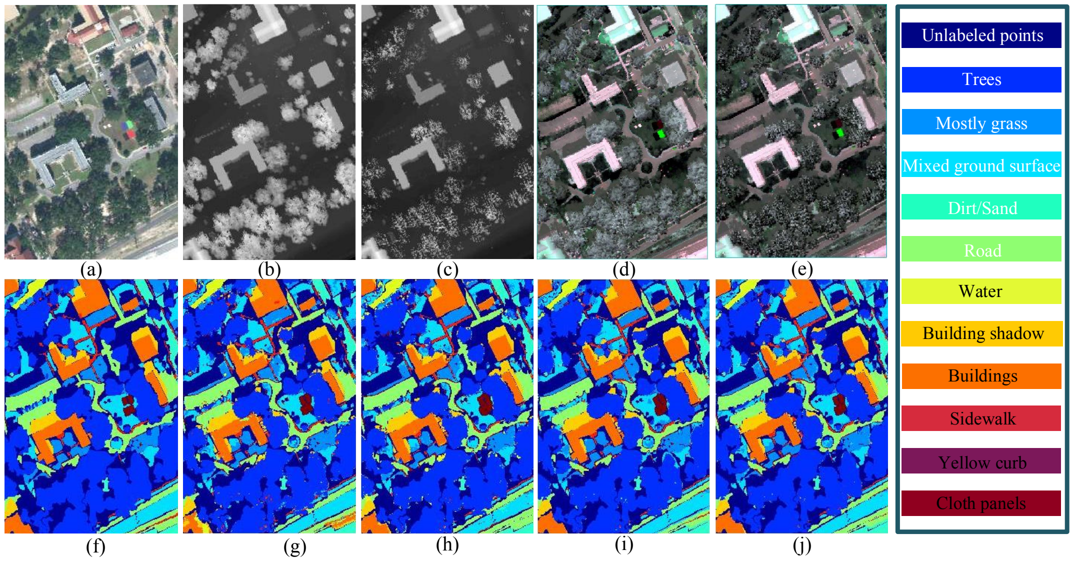

Fig. 3. Thematic maps resulting from classication for the MUUFL Gulfport data with 11 classes: (a) False-color image, (b) LiDAR1 image, (c) LiDAR2 image, (d) Information fusion image1, (e) Information fusion image2, (f) Ground-truth map, (g) Two-branch CNN(H): 88.90%, (h) Two-branch CNN(H+L2): 89.79%, (i) IF-CNN(H+L1): 90.64%, (j) IF-CNN(H+L2): 92.51%.

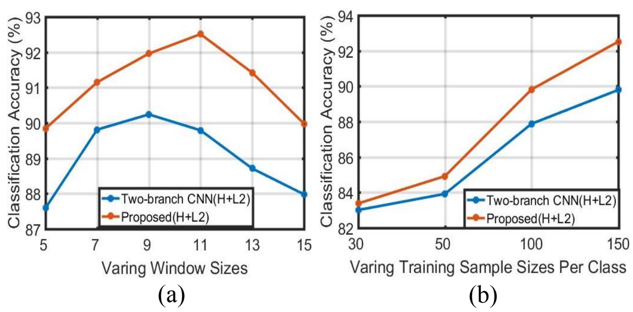

Fig. 4. Overall classification accuracy (%) of various bolck window sizes and different numbers of training samples per class.

after information fusion. Experimental results demonstrate that the proposed IF-CNN can provide statistically better performance than state-of-the-art classifers.

5. REFERENCES

[1] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on geoscience and remote sensing*, vol. 42, no. 8, pp. 1778–1790, Aug 2004.

[2] W. Li, C. Chen, H. Su, and Q. Du, "Local binary patterns and extreme learning machine for hyperspectral imagery classification." *IEEE Trans. Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3681–3693, Jul 2015.

[3] F. A. Mianji and Y. Zhang, "Robust hyperspectral classification using relevance vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2100–2112, Jun 2011.

[4] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 809–823, Mar 2012.

[5] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, Jun 2016.

[6] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *Journal of Sensors*, vol. 2015, Jan 2015.

[7] M. Zhang, W. Li, and Q. Du, "Diverse region-based cnn for hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2623–2634, Jun 2018.

[8] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 844–853, Feb 2017.

[9] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 937–949, Feb 2018.