2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications

CapsNet and Triple-GANs Towards Hyperspectral Classification

Xue Wang
Jiangsu Key Laboratory of Resources and Environment Information Engineering
China University of Mining and Technology,
Xuzhou, China
wx_cumt@yeah.net

Kun Tan
Jiangsu Key Laboratory of Resources and Environment Information Engineering
China University of Mining and Technology,
Xuzhou, China
tankuncu@gmail.com

Yu Chen
Jiangsu Key Laboratory of Resources and Environment Information Engineering
China University of Mining and Technology,
Xuzhou, China
chenyurs@gmail.com

***Abstract*—Hyperspectral processing technology becomes one of the most focused issues in remote sensing field. In the hyperspectral classification, significant improvements have been achieved by various deep learning methods. In general, deep learning algorithms adopt a cascade of layers to extract the hierarchical features. However, the deep hierarchical property will cause some defects such as overfitting and gradient vanishing. In this paper, a hybrid method based on CapsNet and Triple-GANs has been explored to avoid overfitting and extract the effective features. Unlike ordinary CNN, the CapsNet is the consist of a group of capsules with vectorizing the activation output which could consider not only spectral deep features but also the relative locations of these features. The Triple-GANs is a game system with three players: a generator, a classifier and a discriminator. When the Triple-GANs converges to balance, the credible labelled samples could been obtained by the generator which boost the CapsNet in the classification task.**

The main contents are as follows: 1)By introducing the CapsNet in the hyperspectral feature extraction, the 2D convolution operations are replaced by 1D to adapt the pixel-wised spectral features. 2) Use the Triple-GANs and CapsNet to do the hyperspectral classification on small training dataset. Experimental results show that this algorithm can obviously improve the performance of classification compared with the traditional methods.

Keywords*— *CapsNet, Triple-GANs, hyperspectral classification

I. INTRODUCTION

Recently, more and more satellites are launched with hyperspectral sensors, which can produce hyperspectral images data with rich spectral information. As a result, many researchers have been keeping on the investigating hyperspectral techniques in a variety of fields, such as agriculture, forest, monitoring water and mapping etc. Generally, the existing methods for hyperspectral image classification are just at the conventional pattern recognition, such as multinomial logistic regression, artificial neural network, extreme learning machine and support vector machine etc. Deep learning is the latest research in the field of artificial intelligence, which has brought new opportunities for hyperspectral techniques. Some researchers have adopted the deep learning network to simultaneous optimization of feature extraction and classification stages in hyperspectral images, which shows superior performance than the traditional algorithms in machine learning[1-2]. The deep networks used in hyperspectral images are divide into discriminative model such as convolution neural network (CNN), recurrent neural network (RNN) and stacked denoising autoencoder (SAE), and generative model such like deep belief nets (DBNs), deep restricted Boltzmann machine (DRBM), generative adversarial networks (GANs). Mingyi He[3] used 3D CNN for hyperspectral image classification in mult-scale; Weiwei Song[4] proposed a deep feature fusion network to optimize the general CNN in hyperspectral classification; Haokui Zhang[5] adopted one-dimension kernels to fit the hyperspectral context; Tan [6]introduced DBNs into the active learning framework to improve the hyperspectral classification performance; Lichao Mou[7] used RNN model to analyze hyperspectral pixels as sequential data to do the classification task; Zhang Jing[8] used deep convolutional GANs for feature extraction in retrieval task.

However, the conventional CNN has three issues to fit the hyperspectral classification:1) convolution in spatial dimension will mix produce classification error with local pixels and the spatial-spectral convolution operations will cause lots of similar results. 2) Many local features will be lost by max pooling operation. 3) huge amount parameters need significant amount of train datasets while the availability of labelled samples is limited in the scenario of hyperspectral image classification is limited.

This paper aims at overcoming the deficiencies of exist CNN in the field of hyperspectral classification. In this regard, we explore the CapsNet[9] and modify the original architecture to fit the hyperspectral techniques. The most important difference with from the convectional CNN is that the CapsNet has capsules which consist of a group neurons from and conventional CNN is the activity vector. The capsule's activity vector represents the instantiation parameters of a specific type of entity. The length of the activity vector represents the probability that the entity exists and its orientation to represents the instantiation parameters, in other words, the activities of the neurons within active capsules represent the various properties of one particular feature that is present in the hyperspectral data.

978-1-5386-6642-5/18/$31.00 ©2018 IEEE

These properties are potential high dimensional characters characteristics which obtained by convolution filtering. It may be noted here that, in this research, GANs was is used to generate enough lablelled samples to ensure the CapsNet could been trained adequately. GANs have shown promise in image generation and semi-supervised learning (SSL). However, the original GANs in hyperspectral classification have two problems: 1) the generator and the classifier may not be optimal at the same time; 2) the generator cannot control the variance of different categories of the generated samples. Hence, Triple-GANs[10], which is composed of three networks named generator, classifier and discriminator respectively, is introduced to address above-mentioned two issues.

II. Methods

The CapsNet we have explored in this work is modified with 1D convolution to fit the pixel-wised hyperspectral classification which presented in Fig.1.

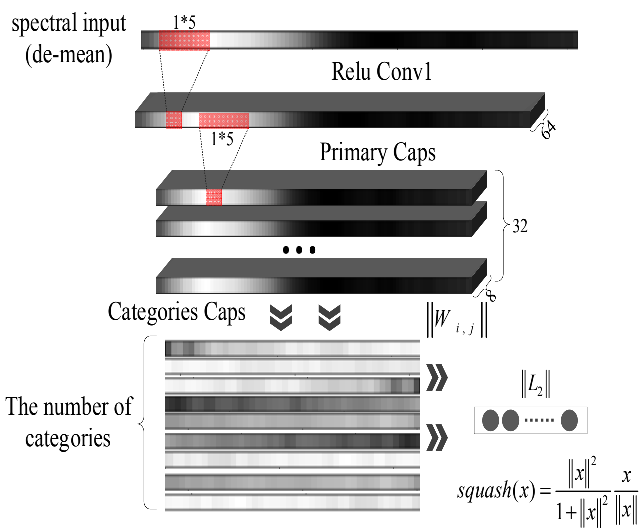


Fig. 1 The structure of CapsNet towards Hyperspectral images

The spectral input is convoluted by the 1D filter at first, and Relu is chosen as activation function. The spectral features derive at different hierarchies are appended to obtain the primary capsules. Each primary capsule is connected to each of the spectral-class capsule, and the predictions are obtained through dynamic routing. The length of the output vector of a capsule would represent the probability that the entity represented by the capsule is present in the current input. Therefor a non-linear "squashing" function is used to ensure that the vectors get shrunk to 0~1.

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|} \quad (1)$$

Where v_j is the vector output of capsule j and s_j is its total input. To allow for multiple classes, we use a separate margin loss L_k for each category capsule, k:

$$L_k = T_k \max\left(0, m^+ - \|v_k\|\right)^2 + \lambda(1-T_k)\max\left(0, \|v_k - m^-\|\right)^2 \quad (2)$$

Where $T_k=1$ if and only if a class k is present, $m^+=0.9$ and $m^-=0.1$. The down-weighting factor λ of the loss for absent classes stop the initial learning from shrinking the lengths of the activity vectors of all the category capsules, we use $\lambda=0.5$. The total loss is simply the sum of the losses of all category capsules.

Moreover, Triple-GANs is explored to generate the abundant train data for improving the performance of the CapsNet. Triple-GANs consists of three players: a generator, a discriminator and a classifier. The generator characterizes the conditional distribution between the spectral data and its label, the discriminator focuses on identifying fake samples, the classifier uses CNN to obtain the posterior probability of each labelled data. The game with three players is designed to ensure the distributions learned by the classifier and the generator both converge to the true data distribution, in other words, the generator could be utilized to obtain labelled samples when the game converge to Nash equilibrium.

The architecture of Triple-GANs in this study consists of three components:1) a classifier C that characterizes the conditional distribution $p_c(y|x) \approx p(y|x)$;2)a class-conditional generator G that characterizes the distribution $p_g(x|y) \approx p(x|y)$;3)a discriminator D that distinguishes whether a pair of data (x,y) come from the true distributionsss $p(x,y)$.Then, the utilities in the process as adversarial losses are used, which could be formulated as a minimax game:

$$\min_{C,G}\max_{D} U(C,G,D) = E_{(x,y)\sim p(x,y)}[\log D(x,y)] + \alpha E_{(x,y)\sim p_c(x,y)}\left[\log\left(1-D(x,y)\right)\right] + (1-\alpha)E_{(x,y)\sim p_g(x,y)}\left[\log\left(1-D(G(y,z),y)\right)\right] \quad (3)$$

Where α is a constant that controls the relative importance of G and C, in this study it set to 0.5 for a balance case.

III. Experiment and Analysis

A. Comparison with the state-of-the-art

Two real hyperspectral images were used to evaluate the approaches. The first hyperspectral image was collected by AVIRIS sensor over the Indian Pines, 202 channels were used in the experiment. The other one was collected by ROSIS sensor over the urban area of the University of Pavia, Italy,103 channels were used in the experiment.

The architecture of CapsNet for Pavia university realized as: The input size is 103*1, the kernel size of Relu layer and Primary Capsules layer is 1*5*64, 1*5*32 respectively, the number of weight in the Category Capsules layer is 1*10*95 and the length of output vector is eight ; the architecture for Indian Pine realized as: The input size is 202*1, the kernel size of Relu layer and Primary Capsules layer is 1*9*64, 1*9*32 respectively, the number of weight in the Category Capsules layer is 1*10*194 and the length of output vector is 8. The mini-batch size for training is 80, learning rate is set to 0.01, down-weighing factor (λ) is set to 0.5. Both Kappa statistics and overall accuracy are employed for the classification results analysis.

978-1-5090-1479-8/16/$31.00 ©2018 IEEE

We chose the conventional classifier like SVM and some mainstream deep learning algorithms as contrast experiments. The SVM in this study used Gaussian radial basis function (RBF) and Liner kernel respectively, the SAE and DBNs used 2 layers and CNN was based on 1D convolution operation.

The results show that, the CapsNet got the best result 88.24% in the experiment of Indian pine, each deep learning algorithm got better performance than SVM with two kernel functions. Nevertheless, the DBNs won by a narrow margin,94.60% on Pavia University dataset, the OA of CapsNet was 94.57%.

Table I. CLASSIFICATION RESULTS BY DIFFERENT METHODS ON INDIAN PINES

Algorithm	OA(%)	Kappa
SVM(Linear)	76.97	0.7706
SVM(RBF)	78.73	0.7885
SAE	82.47	0.8004
DBNs	80.79	0.7811
CNN	82.35	0.7935
CapsNet	**88.24**	0.8572

Table II. Classification results by different methods on Pavia University

Algorithm	OA(%)	Kappa
SVM(Linear)	91.28	0.9128
SVM(RBF)	94.20	0.9354
SAE	92.23	0.9092
DBNs	**94.60**	0.9283
CNN	93.78	0.9083
CapsNet	94.57	0.9172

B. Train in small labelled samples

In this experiment, we explored the collections framework with the Triple-GANs and CapsNet on a small training sample. Three experiments were designed on 1%, 5% and 10% of training dataset respectively. At the beginning of the experiment, the Triple-GANs was been trained with the true training data. When the game was converged or training epochs achieved a certain number, the output from generator could been regarded as a reliable pair of data (x, y). Afterwards, the CapsNet was trained by the hybrid of generated data pairs and the real data pairs.

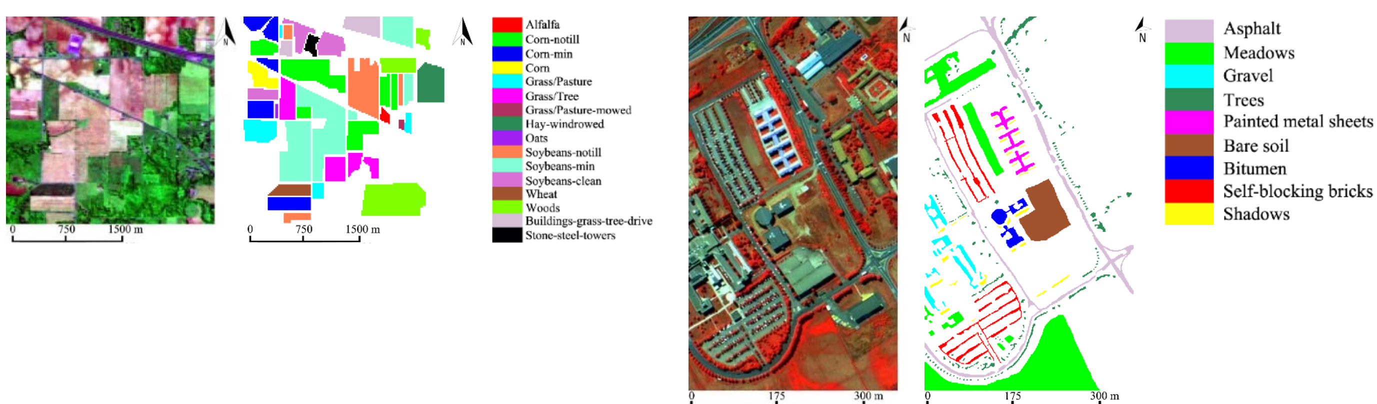
The results shown that: Both CapsNet and Triple-GANs had the worst performance on 1% Indian Pine training data, and the OA were 44.58% and 32.14% respectively. Except on the 1% and 5% Indian Pine training data and 1% Pavia University training data, the Triple-GANs + CapsNet had better performance than Triple-GANs solely. The Triple-GANs in 10% Indian Pine and Pavia University training samples obtained the best performance than other algorithms. The performance of CapsNet depended on the reliability of generator in Triple-GANs. When we use the 1% training data, some categories have just one pair data which causes the Triple-GANs badly training and the generated data poor usability. Whereas, a fully trained Triple-GANs could generate credible labelled samsples that assist in the training of CapsNet to a better performance than its own classifier.

Table III. CLASSIFICATION RESULTS WITH DIFFERENT SIZE TRAINING SAMPLESON INDIAN PINES

Algorithm	OA(%)	Kappa
Triple GAN 0.01	44.58	0.2762
Triple GAN 0.01+CapsNet	32.14	0.2047
Triple GAN 0.05	78.12	0.7513
Triple GAN 0.05+CapsNet	73.75	0.6943
Triple GAN 0.1	87.67	0.84.71
Triple GAN 0.1+CapsNet	**89.70**	0.9317

Table IV. CLASSIFICATION RESULTS WITH DIFFERENT SIZE TRAINING SAMPLES ON PAVIA UNIVERSITY

Algorithm	OA(%)	Kappa
Triple GAN 0.01	83.43	0.7478
Triple GAN 0.01+CapsNet	82.14	0.7240
Triple GAN 0.05	89.43	0.8731
Triple GAN 0.05+CapsNet	93.75	0.8801
Triple GAN 0.1	94.32	0.9097
Triple GAN 0.1+CapsNet	**95.70**	0.9317



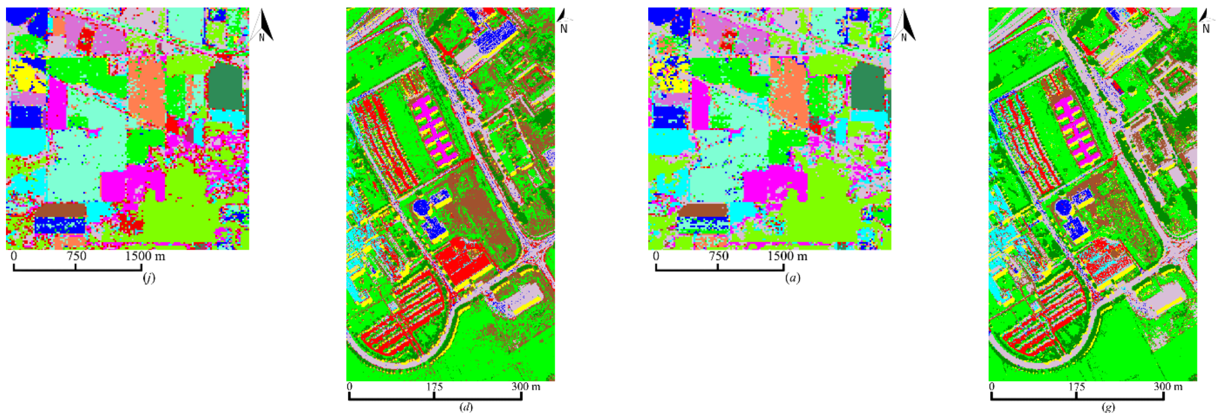
Ground truth test data label
(a) Indian Pines

Ground truth test data label
(b) Pavia University

Fig.2 Two Hyperspectral image

IV. CONCLUSION

This paper explored an innovational network named CapsNet which takes into account the features and their locations as well as their directions. The CapsNet has been modified to adapt to the hyperspectral classification. From the results of two standard hyperspectral datasets, we indicate the validation of the CapsNet and the outperformance compared with other methods. What's more, the Triple-GANs has been adopted to boost the performance of CapsNet on small training data. The results show that unbalanced train dataset among categories affects the training of Triple-GANs and a reliable generator in Triple-GANs could improve the performance of the CapsNet.



(a) CapsNet_IndinaPine_0.1 (b) DBNs_PaviaUniversity_0.1 (c) Triple GAN+CapsNet_IndinaPine0.1 (d) Triple GAN +CapsNet_ PaviaUniversity0.1

Fig.3 The best results of experiments

ACKNOWLEDGMENT

This research is supported in part by Natural Science Foundation of China (No. 41471356), the Xuzhou Scientific Funds (KC16SS092), and Priority Academic Program Development of Jiangsu Higher Education Institutions.

REFERENCES

[1] Chen Y, Lin Z, Zhao X, et al. Deep Learning-Based Classification of Hyperspectral Data[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 7(6):2094-2107.
[2] Chen Y, Zhao X, Jia X. Spectral–Spatial Classification of Hyperspectral Data Based on Deep Belief Network[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2015, 8(6):2381-2392.
[3] He M, Li B, Chen H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]// IEEE International Conference on Image Processing. IEEE, 2017:3904-3908.
[4] Song W, Li S, Fang L, et al. Hyperspectral Image Classification With Deep Feature Fusion Network[J]. IEEE Transactions on Geoscience & Remote Sensing, 2018, PP(99):1-12.
[5] ZHANG H, LI Y. Spectral-spatial classification of hyperspectral imagery based on deep convolutional network; proceedings of the Orange Technologies (ICOT), 2016 International Conference on, F, 2016 [C]. IEEE.
[6] Tan K, Wang X, Zhu J, et al. A novel active learning approach for the classification of hyperspectral imagery using quasi-Newton multinomial logistic regression[J]. International Journal of Remote Sensing, 2018, 39.
[7] Mou L, Ghamisi P, Zhu X X. Deep Recurrent Neural Networks for Hyperspectral Image Classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2017, 55(7):3639-3655.
[8] Zhang J, Chen L, Zhuo L, et al. An Efficient Hyperspectral Image Retrieval Method: Deep Spectral-Spatial Feature Extraction with DCGAN and Dimensionality Reduction Using t-SNE-Based NM Hashing[J]. Remote Sensing, 2018, 10(2):271.
[9] Sabour S, Frosst N, Hinton G E. Dynamic Routing Between Capsules[J]. 2017.
[10] Li C, Xu K, Zhu J, et al. Triple Generative Adversarial Nets[J]. 2017.

978-1-5090-1479-8/16/$31.00 ©2018 IEEE