

A Classification Method for Hyperspectral Imagery Based on Deep Learning

Lin Yuan¹, Shaoxing Hu¹, Aiwu Zhang², Shatuo Chai³, Xing Wang³

¹School of Mechanical Engineering and Automation, Beihang University, Beijing

²College of Resource Environment and Tourism, Capital Normal University, Beijing

³Animal husbandry and Veterinary Hospital of Qinghai University, Xining Qinghai

Email: yuanlinhyhy@163.com

Received: Feb. 3rd, 2017; accepted: Feb. 18th, 2017; published: Feb. 24th, 2017

Abstract

Remote sensing hyperspectral imaging can obtain abundant spectral information, which provides the possibility for the analysis of high precision terrain. The hyperspectral image has the characteristics of "map in one", and the full use of spectral information and spatial information in hyperspectral image is the premise of obtaining accurate classification results. Deep learning stack machine model in automatic encoding (Stack Auto-Encoder SAE) can effectively extract data in nonlinear information, and convolutional neural network (Convolutional Neural Network, CNN) can automatically extract features from the image. Based on this, this paper presents a classification method of hyperspectral images based on deep learning. Firstly, the spectral dimension of the hyperspectral data is reduced using automatic encoding machine, then convolutional neural network is used as the classifier, and the pixel and its neighborhood pixels are classified together as the input of the classifier, so as to realize the hyperspectral image classification with spectral space.

Keywords

Hyperspectral, Image Classification, Depth Learning, Automatic Coding Machine, Convolutional Neural Network

基于深度学习的高光谱图像分类方法

袁林¹, 胡少兴¹, 张爱武², 柴沙陀³, 王兴³

¹北京航空航天大学机械工程及自动化学院, 北京

²首都师范大学资源环境与旅游学院, 北京

³青海大学畜牧兽医院, 青海 西宁

Email: yuanlinhyhy@163.com

收稿日期: 2017年2月3日; 录用日期: 2017年2月18日; 发布日期: 2017年2月24日

摘要

遥感高光谱成像能够获得丰富的地物光谱信息,这使得在传统的宽波段遥感中不可分辨的物质,在高光谱遥感中可以被分辨出来。高光谱图像具有“图谱合一”的特点,充分的利用高光谱图像中的光谱信息和空间信息是获得精确分类结果的前提。深度学习模型中的自编码神经网络能够实现高维数据的非线性降维,而卷积神经网络(Convolutional Neural Network, CNN)则能够自动的从图像中提取空间特征,基于此,本文提出了一种基于深度学习的Autoencoder-CNN高光谱图像分类方法。首先利用自编码神经网络对高光谱数据进行光谱维的降维,然后将卷积神经网络作为分类器,将待分类像元及其邻域像元一同作为卷积神经网络的输入,实现高光谱图像的空谱联合分类。

关键词

高光谱, 图像分类, 深度学习, 自编码神经网络, 卷积神经网络

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

通过结合成像技术和光谱技术,高光谱遥感可以同时获得空间和光谱连续的数据。光谱数据在检测地球表面中是一种有效的工具,被广泛用于农业,矿物学,对地观测,物理学,天文学,化学成像和环境科学中[1]。这些应用中的常用技术是对高光谱图像中每个像素进行分类。但是,高光谱图像的高数据维与不断提高的空间分辨率给传统的分类任务提出了新的挑战,提高高光谱图像的分类精度一直是遥感领域的研究热点。

限制高光谱图像分类精度的主要原因有两个:第一,与传统的图像分类问题[2]相比,高光谱图像的分类问题具有数据空间维数高,训练样本难以获得的特点,即所谓的“维数灾难”[3],这使得传统的在低维空间中表现良好的分类方法在高光谱图像分类问题上表现的不尽如人意。传统的解决维数灾难的方法是对高光谱数据进行降维处理后再分类[4][5][6],但传统的线性特征提取方法,如PCA、ICA、NFWE等,在降维的过程中舍弃了高光谱图像中光谱维上的细节信息,降维后的图像与多光谱无异,因而失去了“高光谱”的意义。第二,传统的方法把高光谱图像中的像元作为独立的光谱曲线进行分类,忽视了图像中的空间信息[6]。为了利用图像中的空间信息,已有的文献从图像中提取纹理特征,结构特征,形态学特征等空间特征作为光谱特征的补充[6][7][8]。但是各种的空间特征的抽取都需要人员的干预甚至设计,最终的分类效果与特征的好坏具备直接的关系,导致分类的准确率很大的依赖于人的经验,可能某一种空间特征在某一数据集上表现良好但是在其他数据集上结果却完全相反。

综上所述,高光谱图像分类问题在降维和空间特征提取两方面均有改进的空间,因此,本文提出了一种基于深度学习的高光谱图像分类方法,针对已有的分类方法中的缺陷,关注了深度学习中的两种常见模型——自动编码网络和卷积神经网络。自编码网络[9]可以实现非监督的提取数据的特征,恰当设置非线性激活函数即可有效提取得到数据的非线性特征。卷积神经网络[10][11]作为一种成功的视觉模型,已经在图像分类[2],人脸识别[12],目标检测[13]等领域取得了巨大的成功,被证明能够自动从图像中有

效的提取对分类结果有益的特征[14]，避免了人工设计并抽取特征的过程。本文的分类算法流程图如图1所示。首先利用堆叠自动编码器对高光谱数据进行光谱维上的降维，然后将卷积神经网络作为分类器，将以待分类像元为中心的矩形内的高光谱数据立方体作为卷积神经网络的输入，进行空谱联合分类，获得最终的分类结果。

2. 基于堆叠自编码网络的高光谱数据降维

自动编码网络由其基本单位——自动编码器堆叠而成。自动编码器是一个三层前馈神经网络，由输入层，隐藏层，重构层构成。自动编码器的编码及解码过程如公式[15] (1) (2):

$$y = \delta(W_y x + b_y) \quad (1)$$

$$z = \delta(W_z y + b_z) \quad (2)$$

式中， y 是原始数据的特征表达， W_y ， W_z ， b_y ， b_z 分别为输入层到隐藏层，隐藏层到重构层的权重及偏移系数，一般取 $W_y = W = W_z'$ ， $\delta(\cdot)$ 为非线性映射函数，一般取为 sigmoid 函数，即：

$$\delta(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

通过调节 W ， b_y ， b_z 使得输入 x 和重构 z 之间相似。一般采用交叉熵函数衡量 x 和 z 之间的距离，训练时采用分批训练，每一批的样本数为 m ，则损失函数为：

$$L(x, z) = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^d [x_{ik} \log(z_{ik}) + (1 - x_{ik}) \log(1 - z_{ik})] \quad (4)$$

公式(4)中， m 是每一批训练样本的个数， d 是输入层与重构层的维数。

使用随机梯度下降训练网络参数，参数的更新规则为(η 代表学习率):

$$W = W - \eta \frac{\partial L(x, z)}{\partial W} \quad (5)$$

$$b_y = b_y - \eta \frac{\partial L(x, z)}{\partial b_y} \quad (6)$$

$$b_z = b_z - \eta \frac{\partial L(x, z)}{\partial b_z} \quad (7)$$

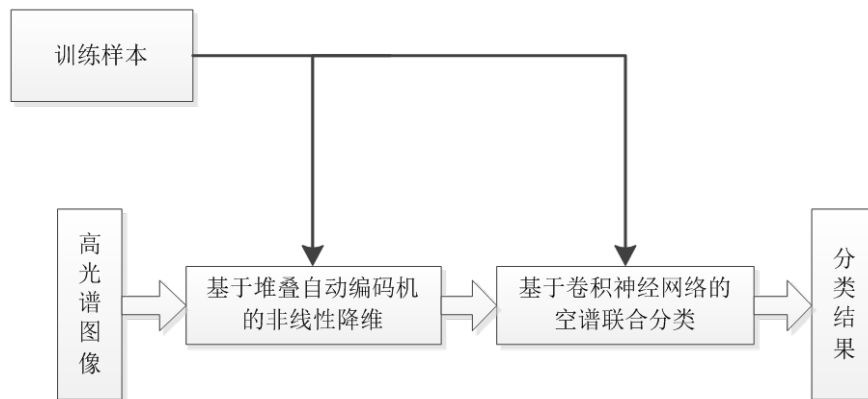


Figure 1. The flowchart of the proposed algorithm

图1. 本文分类流程

自动编码网络由若干个自动编码器层叠而成，上一层自动编码器隐藏层的输出作为下一层自动编码器的输出层的输入。其训练共分为三个阶段[15]，共包括预训练，展开和微调 3 个步骤：

步骤 1 预训练过程。逐层训练组成自动编码网络的多个自动编码器，下层自动编码器隐藏层单元输出作为其上层自动编码器的输入参与训练。

步骤 2 展开过程。预训练完成后，下层 AE 输出单元与其上层 AE 合并为一层，将多个 AE 连接成一个自动编码深度网络。

步骤 3 微调过程。展开的自动编码网络采用反向传播算法对预训练得到的初始权值进一步调整，进一步减少误差。

3. 基于卷积神经网络的空谱联合分类

卷积神经网络由卷积层，池化层，全连接层和 softmax 分类层组成[15]。一般地，在卷积层中对网络的输入或者前一个隐藏层的输出进行卷积操作生成特征图，卷积操作生成的每一张特征图都会与偏置项 b_l 相加，随后非线性激活函数会作用在特征图中的每一个像素上。接下来，池化层会以非重叠的方式从每一个特征图中选取局部的主特征，也就是对特征图进行降维操作。整个过程可以用公式[16]表示为：

$$H^l = \text{pool}\left(g\left(H^{l-1} * W_l + b_l\right)\right) \quad (8)$$

式中，*代表卷积操作，pool 代表最大化池操作。

数个卷积层和池化层交替重叠能够从输入中逐层的提取特征。之后，提取的特征在拉伸成为一个向量后送入全连接层。在全连接层，首先通过与权重矩阵 W_l 相乘和偏置 b_l 相加进行线性变换，然后将非线性函数作用在变换后的每一个分量上，有：

$$H^l = g\left(H^{l-1} * W_l + b_l\right) \quad (9)$$

激活函数 $g(n)$ 取公式(3)中的 sigmoid 函数。全连接层的输出被送至 softmax 层进行分类。卷积神经网络中的参数为权重矩阵 $\{W_1, \dots, W_L\}$ 以及偏置 $\{b_1, \dots, b_L\}$ 。网络通过反向传播算法进行训练，损失函数为：

$$\min L = \frac{1}{2} \sum_{i=1}^M \|t_i - y_i\|_2^2 \quad (10)$$

通过使用随机梯度下降[11]算法获得最优的参数值 W 和 b 。

为了对降维后的高光谱数据进行空谱联合的分类，本文设计了如图 2 所示的卷积神经网络结构。取以待分类像元为中心的矩形内的数据立方体作为卷积神经网络的输入，取矩形的大小为 7×7 ，则网络输入的尺寸为 $7 \times 7 \times N$ ，其中 N 是高光谱图像降维后的光谱维度。与一般的图像分类方法，如人脸识别相比，本文分类任务的空间维度并不高，因此本文的卷积神经网络中不包含池化层的空间降维操作。卷积神经网络由三个卷积层，一个全连接层和一个 softmax 分类层组成。每个卷积层的卷积核大小为 3×3 ，卷积核的个数为 N ，全连接层的输入维数均为 N ，输出维数为 30，softmax 分类层的输入个数为 N ，输出个数为高光谱图像中的地物种类数。

4. 实验和分析

4.1. 实验数据集描述

本实验使用两组高光谱数据集检验本文算法的分类性能，分别是 IndianPines 数据集和 University of Pavia 数据集。前者为 AVIRIS 采集的农业区高光谱图像，图像大小为 145 像素 \times 145 像素，共包含 220 个波段，去掉其中的 20 个水吸收严重的波段，得到包含 200 个波段的高光谱数据。后者为 ROSIS 采集

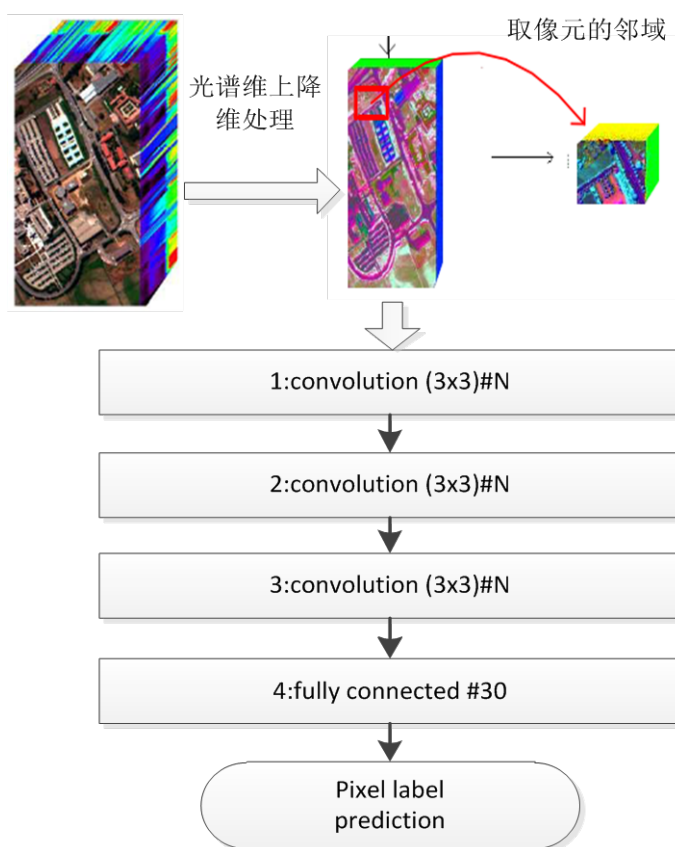


Figure 2. The structure of convolution natural network for hyperspectral classification
图 2. 用于高光谱图像分类的卷积神经网络结构

的城区高光谱图像，图像大小为 610 像素 × 340 像素，共包含 115 个波段，去除其中水吸收严重的 12 个波段，将其余 103 个波段作为高光谱数据。Indian Pines 数据以及 University of Pavia 数据的假彩色图及标记模板如图 3(a)、图 3(b)和图 4(a)、图 4(b)所示。

4.2. 实验结果与讨论

4.2.1. Indian Pines 数据集

为了证实本文的算法确实有效，将本文的算法与以下分类算法进行对比：

1) PCA-SVM [9]：利用主成分分析对高光谱图像进行降维，然后采用 SVM 方法进行分类。

2) Autoencoder-SVM [15]：利用自编码网络对高光谱进行降维，然后使用 SVM 分类器进行分类，此结果用于证明 Autoencoder 在降维方面的优越性。

3) PCA-MOR-SVM [9]：利用主成分分析对高光谱数据进行降维，然后对前 3 个主成分提取形态学特征作为光谱特征的补充[1]，接下来采用 SVM 分类器进行分类

4) PCA-CNN：利用主成分分析对高光谱数据降维，然后使用本文提出的卷积神经网络进行分类。

PCA 和 Autoencoder 在降维方面各有优势，PCA 能够用较少的维数表示原始数据中的大部分信息，而 SAE 则能够更好的保留数据的特征。本文的堆叠自动编码机的结构参照文献[14]设置，经过堆栈自动编码器降维后的数据维数为 40，自动编码器具有 4 层结构，每一层的节点数为 200-40-40-40。在 1)中同样取前 40 个主成分进行分类；在 2)中利用 SAE 降维后的维数为 40，并且使用 SVM 进行分类；在 3)中

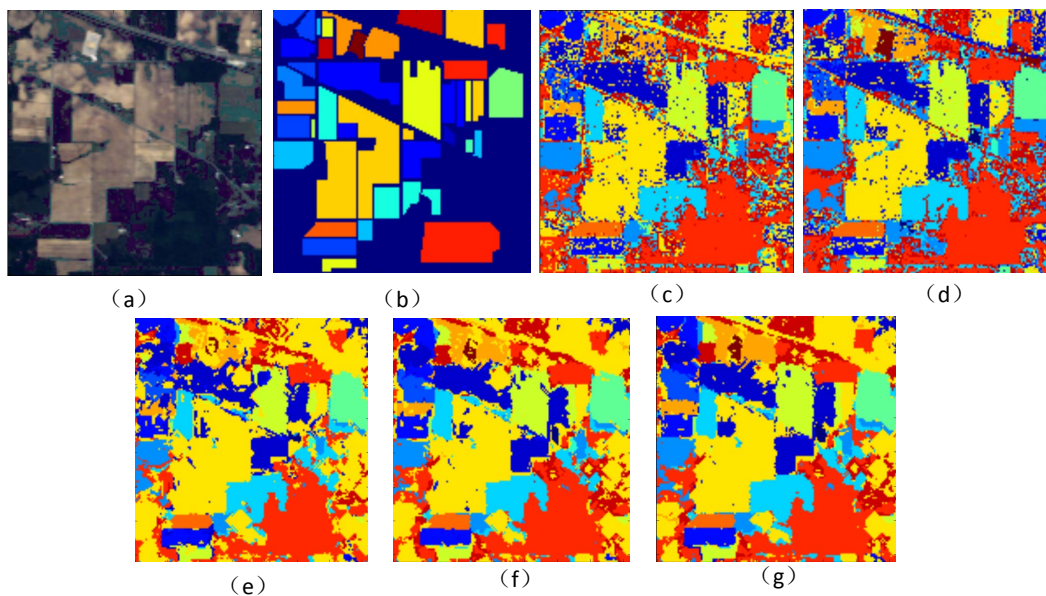


Figure 3. The result of classification of Indian Pines. (a) Pseudocolor image; (b) Ground truth map; (c) PCA-SVM: 80.47%; (d) Autoencoder-SVM: 85.48%; (e) PCA-MOR-SVM: 91.13%; (f) PCA-CNN: 95.27%; (g) Autoencoder-CNN: 98.64%

图3. Indian Pines数据集分类结果。(a) 假彩色图像; (b) 分类参考图; (c) PCA-SVM: 80.47%; (d) Autoencoder-SVM: 85.48%; (e) PCA-MOR-SVM: 91.13%; (f) PCA-CNN: 95.27%; (g) Autoencoder-CNN: 98.64%

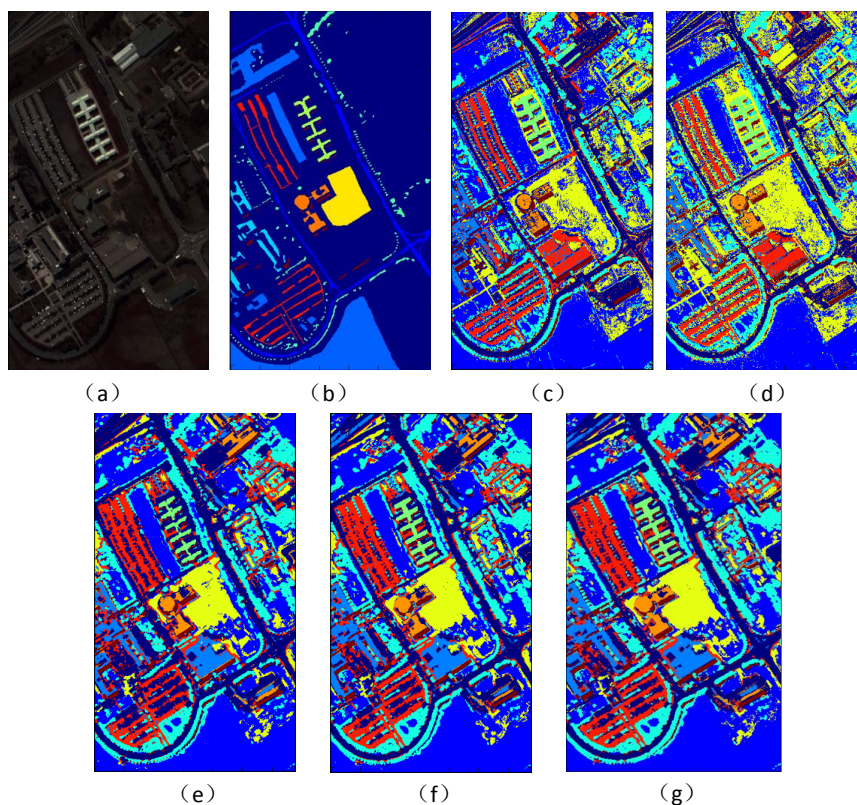


Figure 4. The classification result of University of Pavia. (a) Pseudocolor image; (b) Ground truth map; (c) SVM: 80.01%; (d) Autoencoder-SVM: 94.14%; (e) PCA-MOR-SVM: 98.16%; (f) PCA-CNN: 98.56%; (g) Autoencoder-CNN: 99.26%

图4. Indian Pines数据集分类结果。(a) 假彩色图像; (b) 分类参考图; (c) SVM: 80.01%; (d) Autoencoder-SVM: 94.14%; (e) PCA-MOR-SVM: 98.16%; (f) PCA-CNN: 98.56%; (g) Autoencoder-CNN: 99.26%

同样取前 40 个主成分进行分类, 并且依据文献建立半径为 2 的圆形结构元素, 对前 3 个主成分进行四次开闭运算, 得到的形态学特征作为光谱特征的补充; 在 4) 取 PCA 降维后的前 40 个主成分, 并使用本文提出的 CNN 进行空谱联合分类。表 1 给出了图像详细的训练训练集的个数, 测试集的个数以及分类数量。图 3 中为一次实验的结果图, 图 5 是每一种地物的分类情况。从视觉角度看, 仅利用光谱信息的分类对噪声比较敏感, 分类结果中含有许许多多细小斑点, 而综合利用光谱与空间信息的方法则能够获得比较完整的地物区域。

4.2.2. University of Pavia 数据集

为了进一步证明本文的算法确实有效, 使用同样的方法对 University of Pavia 数据集进行分类。PCA 降维方式同样保留前 40 个主成分, Autoencoder 的结构为 103-40-40-40, 即使用深度为 4 的自动编码网络进行降维处理, 降维后的光谱维度为 40。Pavia 大学数据集的类别数量以及训练和测试样本数量如表 2 所示, 图 4 为分类的结果, 图 6 为每一类地物分类的精度。从 Pavia 大学数据的分类结果可以得到与

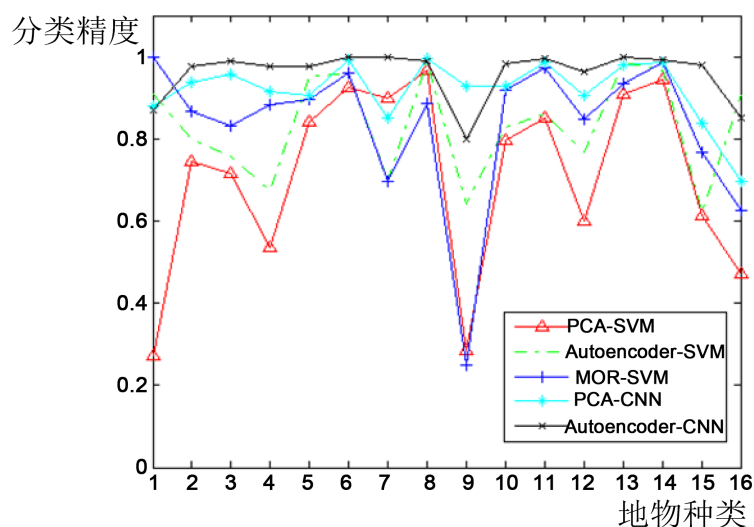


Figure 5. Classification accuracy of Indian Pines data

图 5. Indian Pines 数据分类精度

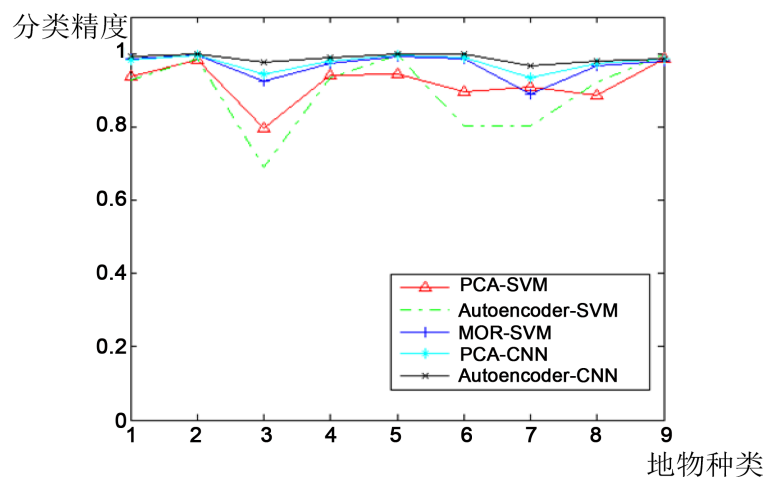


Figure 6. Classification accuracy of Pavia University data

图 6. Pavia University 数据分类精度

Table 1. Class labels and train-test distribution of samples for the Indian Pines datasets**表1.** Indian Pines高光谱影像的类别和样本数

No	Name	Train	Test
1	Asphalt	10	36
2	Corn-notill	428	1000
3	Corn-minitill	200	630
4	Corn	47	197
5	Grass-pasture	150	333
6	Grass-tress	210	520
7	Grass-pasture-mowed	7	19
8	Hay-windrowed	119	360
9	Oats	6	14
10	Soybean-notill	300	672
11	Soybean-minitill	520	1935
12	Soybean-clean	100	493
13	Wheat	50	155
14	Woods	200	1065
15	Build-Grass-Tree-Drives	186	200
16	Stone-Steel-Towers	30	63

Table 2. Class labels and train-test distribution of samples for the University of Pavia**表2.** Indian大学高光谱影像的类别和样本数

No	Name	Train	Test
1	Asphalt	2231	4400
2	Meadows	6993	12,493
3	Gravel	609	1440
4	Trees	1099	2064
5	Painted Metal Sheets	401	945
6	Bare Soil	1500	4029
7	Bitumen	300	1030
8	Self-Blocking Bricks	1000	2682
9	Shadows	300	647

之前的类似结论，即：1) 非线性降维方式优于线性降维方式；2) 加入空间信息有利于分类精度的提升；3) 使用卷积神经网络能够自动的提取对分类有利的特征。

5. 结论

本文从限制高光谱图像分类精度的两点出发，提出了基于深度学习的高光谱图像分类方法。在降维方面，针对 PCA 等线性降维方式不能提取光谱维上的非线性特征这一问题，使用了自编码网络提取光谱

维度的非线性信息；另外，设计了一种卷积神经网络结构，能够自动的提取对分类有利的空间特征，并在分类时能同理利用光谱信息和空间信息；从而获得比传统分类方法更精确的分类结果。

参考文献 (References)

- [1] Landgrebe, D. (2002) Hyperspectral Image Data Analysis. *Signal Processing Magazine*, **19**, 17-28. <https://doi.org/10.1109/79.974718>
- [2] Wange, A., Lu, J., Cai, J., Wang, G. and Charm, T.J. (2015) Unsupervised Joint Feature Learning and Encoding for RGB-D Scene Labeling. *IEEE Transactions on Image Process*, **24**, 4459-4473. <https://doi.org/10.1109/TIP.2015.2465133>
- [3] Huges, G. (1968) On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Transactions on Information Theory*, **14**, 55-63. <https://doi.org/10.1109/TIT.1968.1054102>
- [4] Kuo, B.-C. and Landgrebe, D.A. (2004) Nonparametric Weighted Feature Extraction for Classification for Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 1096-1105. <https://doi.org/10.1109/TGRS.2004.825578>
- [5] Landgrebe, D.A. (2003) *Signal Theory Methods in Multispectral Remote Sensing*. Wiley, Hoboken.
- [6] Liu, Y., Gu, Y. and Zhang, Y. (2006) Hyperspectral Feature Extraction Using Selective PCA Based on Genetic Algorithm with Subgroups. *1st International Conference on Innovative Computing, Information and Control*, Beijing, 30 August-1 September 2006, 652-656.
- [7] Zhao, Y., Zhang, L., Li, P. and Huang, B. (2007) Classification of High Spatial Resolution Imagery Using Improved Gaussian Markov Random-Field-Based Texture Features. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 1458-1468. <https://doi.org/10.1109/TGRS.2007.892602>
- [8] Xu, H., Li, P. and Shen, Y. (2008) The Application of Invariant Moments to High Resolution Remote Sensing Image Classification. *Remote Sensing for Land & Resources*, **20**, 9-13.
- [9] Luo, B. and Zhang, L. (2014) Robust Autodual Morphological Profiles for the Classification of High-Resolution Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing*, **52**, 1451-1462. <https://doi.org/10.1109/TGRS.2013.2251468>
- [10] Bengio, Y., Lamblin, P., Popovici, D., *et al.* (2007) Greedy Layer-Wise Training of Deep Networks. *Advances in Neural Information Processing Systems*, **19**, 153.
- [11] Fukushima, K. (1988) Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition. *Neural Networks*, **1**, 119-130. [https://doi.org/10.1016/0893-6080\(88\)90014-7](https://doi.org/10.1016/0893-6080(88)90014-7)
- [12] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2323. <https://doi.org/10.1109/5.726791>
- [13] Taigman, Y., Yang, M., Ranzato, M. and Wolf, L. (2014) DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 1701-1708.
- [14] Sainath, T.N., Mohamed, A.-R., Kingsbury, B. and Ramabhadran, B. (2013) Deep Convolutional Neural Networks for LVCSR. *Proceedings of the 38th IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vancouver, 26-31 May 2013, 8614-8618. <https://doi.org/10.1109/icassp.2013.6639347>
- [15] Chen, Y., Lin, Z., Zhao, X., Wang, G. and Gu, Y. (2014) Deep Learning-Based Classification of Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7**, 2094-2107. <https://doi.org/10.1109/JSTARS.2014.2329330>
- [16] Lawrence, S., Giles, C.L., Tsoi, A.C. and Back, A.D. (1997) Face Recognition: A Convolutional Neural-Network Approach. *IEEE Transactions on Neural Networks*, **8**, 98-113. <https://doi.org/10.1109/72.554195>

期刊投稿者将享受如下服务：

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：airr@hanspub.org