

结合最小噪声分离变换和卷积神经网络的高分辨率影像分类方法

陈洋^{1,2}, 范荣双², 王竞雪¹, 吴增林¹, 孙汝星¹

¹辽宁工程技术大学测绘与地理科学学院, 辽宁 阜新 123000;

²中国测绘科学研究院, 北京 100830

摘要 针对传统浅层机器学习方法应用于高分辨影像分类时存在的问题,提出了结合最小噪声分离变换和卷积神经网络的高分辨率影像分类方法。采用最小噪声分离分析非监督训练初始化卷积神经网络,为提高训练速度,使用线性修正函数作为神经网络的激活函数;利用概率最大化采样原则减少池化过程中影像特征的缺失,并将下采样后影像特征输入 Softmax 分类器进行分类。采用所提分类方法对典型地区的影像进行分类实验,并与支持向量机和人工神经网络分类方法的分类结果进行对比。结果表明,所提分类方法的分类精度明显高于另两种分类方法的分类精度,并能充分挖掘高分辨遥感影像的空间信息。

关键词 遥感; 高分辨影像; 卷积神经网络; 最小噪声分离变换; 影像分类

中图分类号 P237 文献标识码 A

doi: 10.3788/LOP54.102801

High Resolution Image Classification Method Combining with Minimum Noise Fraction Rotation and Convolution Neural Network

Chen Yang^{1,2}, Fan Rongshuang², Wang Jingxue¹, Wu Zenglin¹, Sun Ruxing¹

¹ School of Geomatics, Liaoning Technical University, Fuxin, Liaoning 123000, China;

² Chinese Academy of Surveying and Mapping, Beijing 100830, China

Abstract Aiming at the problems of traditional shallow machine learning methods applied to high resolution image classification, we propose a high resolution image classification method combining with minimum noise fraction (MNF) rotation and convolution neural networks (CNN). MNF is used to analyze the initial unsupervised pre-training CNN. Linear correction function is adopted as the activation function of the neural network to increase the training speed. In order to reduce the missing of image features in the process of the pool, the sampled image features are put into Softmax classifier under the principle of maximizing sampling probability. Experimental image of typical regions is selected and classified by using the proposed classification method, and the classification results are compared with those of support vector machines classification method and artificial neural network classification method. The results show that the classification accuracy of the proposed method is superior to the shallow machine learning classification methods, and can fully excavate the spatial information of high resolution remote sensing images.

Key words remote sensing; high resolution image; convolution neural network; minimum noise fraction rotation; image classification

OCIS codes 280.4788; 100.2960; 100.2980; 100.3008

收稿日期: 2017-05-23; 收到修改稿日期: 2017-06-06

基金项目: 国家重点研发计划(2016YFC0803100)、国家自然科学基金(41101452)、高等学校博士学科点专项科研基金(20112121120003)、辽宁省教育厅科研项目(LJYL010)

作者简介: 陈洋(1991—),男,硕士研究生,主要从事影像分割、地物信息智能提取方面的研究。

E-mail: 874153187@qq.com

导师简介: 范荣双(1975—),男,博士,研究员,主要从事遥感与地理信息技术应用方面的研究。

E-mail: 13621051775@163.com

1 引言

近年来,随着遥感技术的进步,遥感卫星获取影像的分辨率不断提高。由于高分辨遥感影像中的地物细节突出,类内的光谱差异较大,类间的光谱差异小,从而使得同物异谱和同谱异物的现象更加明显,因此导致基于传统光谱特征统计分类方法(如 K -mean 聚类、极大似然分类和决策树分类等方法)的分类精度较低^[1-2]。在这样的背景下,许多研究人员将支持向量机(SVM)、人工神经网络(NN)等的机器学习算法应用于高分辨影像的分类,提高了分类精度^[3]。但 SVM 和 NN 分类方法都属于浅层机器学习算法,这些浅层结构模型因为计算单元有限,不能充分学习复杂样本中的复杂结构信息^[4-5]。深度学习算法由多层非线性映射层组成,具有学习复杂训练样本的能力,对复杂分类具有很好的稳健性^[6-7]。在高分辨影像地物复杂的背景下,浅层学习算法已不再适用于高分辨影像的高精度分类,深层次的结构模型因包含复杂的多层非线性变换而具有更强的表达与建模能力,更适合用于处理复杂的地物分类。文献[8]提出了基于大型深层卷积神经网络(CNN)的自然图像分类算法,在 ImageNet 数据集上获得了很高的分类精度,该算法采用监督训练的方式,而网络权值的调整需要大量训练样本,若训练样本较少,就会导致卷积神经网络模型不能被训练。最小噪声分离(MNF)变换能充分挖掘高分辨遥感影像的特征(相当于对影像做 2 次主成分变换),通常用于遥感影像特征的提取^[9]。

针对上述高分辨遥感影像分类存在的问题,本文将卷积神经网络算法应用于高分辨遥感影像分类中,提出了结合最小噪声分离和卷积神经网络的高分辨率影像分类方法。首先,采用最小噪声分离分析非监督训练初始化卷积神经网络,在最小噪声分离分析最小化重构误差的原则下,获取了待分类影像的特征,从而得到含有训练样本统计特性的滤波器。为了提高训练速度,本文采用线性修正函数作为神经网络的激活函数;为了减少池化过程中影像特征的缺失,采用概率最大化采样原则,最后将下采样后的影像特征输入 Softmax 分类器进行分类。

2 原理与方法

2.1 最小噪声分离分析非监督训练

设输入卷积神经网络的影像有 M 景,其大小为 $m \times n$,卷积滤波器大小为 $g_1 \times g_2$,在训练影像的第 i 景中将大小为 $g_1 \times g_2$ 的影像块全部取出,并将其表示成向量的形式,即 $\mathbf{X}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,mm})$,对 \mathbf{X} 进行去均值化处理,得到影像 \mathbf{X}_i 的影像块为 $\bar{\mathbf{X}}_i = (\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,mm})$ 。所以,训练数据的影像块可以表示为

$$\mathbf{X} = (\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n), \mathbf{X} \in \mathbf{R}^{g_1 g_2 \times Mmm}. \quad (1)$$

采用最小噪声分离变换方法最小化重构误差求解特征向量:

$$\begin{cases} \min_{\mathbf{V} \in \mathbf{R}^{g_1 g_2 \times H_1}} \|\mathbf{X} - \mathbf{V}\mathbf{V}^T \mathbf{X}\|^2 \\ \text{s.t. } \mathbf{V}^T \mathbf{V} = \mathbf{I}_H \end{cases}, \quad (2)$$

式中 \mathbf{I}_H 为 $H \times H$ 单位矩阵; \mathbf{V} 为协方差矩阵 $\mathbf{X}\mathbf{X}^T$ 的前 H 个特征向量,它可以表示输入影像块的主要特征。采用最小噪声分离变换学习得到初始化卷积神经网络的滤波器组 \mathbf{W}_h :

$$\mathbf{W}_h = m_{g_1 g_2}(\mathbf{V}_h), h = 1, 2, \dots, H, \quad (3)$$

式中 $m_{g_1 g_2}(\mathbf{V}_h)$ 表示将向量 \mathbf{V} 映射到矩阵 \mathbf{W}_h , \mathbf{V}_h 为影像的第 h 个主要特征。

采用最小噪声分离分析非监督训练可以获取训练数据局部影像块的主要特征,故而可以将最小噪声分离分析训练滤波器看作是一种简单的自动编码器。

2.2 卷积神经网络

卷积神经网络是受生物神经学知识启发,参考其结构原理,并结合人工神经网络而取得的开创性研究成果之一,是一种具有深度学习能力的人工神经网络系统。与传统方法相比,卷积神经网络具有适用性强、特征提取与分类同时进行、泛化能力强、全局优化训练参数少等优点,已成为当前的研究热点^[10-12]。卷积神经网络主要由输入层、卷积层、池化层和全连接层组成,其结构如图 1 所示。

输入影像在卷积层和滤波层中进行卷积运算,且卷积层和滤波层通过激活函数连接^[13-14]。用 \mathbf{G}_i 表示卷

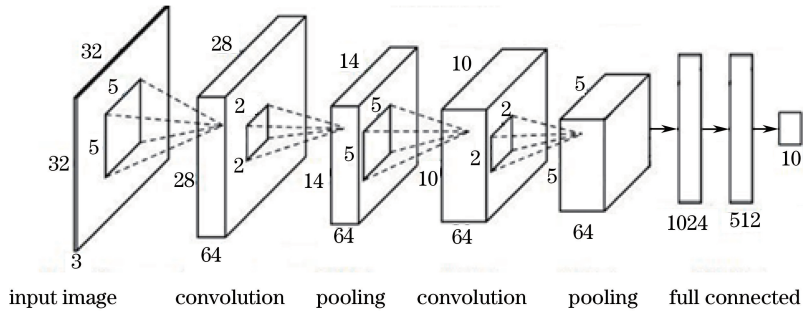


图1 卷积神经网络结构示意图

Fig. 1 Schematic of convolution neural network structure

积神经网络第 i 层特征图,将卷积过程描述为

$$\mathbf{G}_i = f(\mathbf{G}_{i-1} \otimes \mathbf{W}_i + \mathbf{b}_i), \quad (4)$$

式中 \mathbf{W}_i 为第 i 层卷积核的权值向量, \mathbf{b}_i 为偏移向量, $f(\cdot)$ 为线性激活函数, 运算符号 \otimes 表示对第 i 层影像与第 $i-1$ 层影像或特征图进行卷积操作。通过线性激活函数 $f(\cdot)$ 得到第 i 层特征图 \mathbf{G}_i 。

激活函数分为线性激活函数和非线性激活函数。常用的非线性激活函数有双曲线函数、Sigmoid 函数和 Softplus 函数。双曲线函数是 Sigmoid 函数的一种变体, 双曲线函数的取值范围为 $[-1, 1]$, Sigmoid 函数的取值范围为 $[0, 1]$ 。因为线性修正函数与生物神经元受刺激后的激活模型最接近, 且具有一定的稀疏性, 并且计算简单, 所以逐渐被用来作为卷积神经网络的激活函数^[15]。本研究采用 Sigmoid 函数的线性修正函数 $g(x) = \max(x)$ 作为卷积神经网络的激活函数。

对图像进行卷积之后获得特征图, 由于一般中等大小的图像卷积后的特征维度都在百万数量级, 直接训练分类器很容易过拟合, 因此在进行分类前需要对影像的特征图进行下采样, 即池化操作。为减少池化过程中影像特征的缺失, 采用概率最大化进行池化操作。当采样块 \mathbf{X}_n 中有一个神经元为激活状态时, 采样值 P_n 产生响应, 否则不响应。其表达式为

$$P(P_n | i) = \frac{\exp(\mathbf{G}_{i-1} \otimes \mathbf{W}_i + \mathbf{b}_i)}{1 + \exp(\mathbf{G}_{i-1} \otimes \mathbf{W}_i + \mathbf{b}_i)}. \quad (5)$$

结合最小噪声分离变换和卷积神经网络的高分辨率影像分类方法的步骤为: 1) 采用最小噪声分离变换初始化卷积神经网络滤波器的权值; 2) 对影像进行卷积计算得到影像特征图; 3) 对影像特征图采用概率最大化原则进行池化; 4) 将特征图合并成一个列向量, 并输入全连接层, 通过反向传播算法更新滤波器的权值; 5) 利用 Softmax 分类器完成影像分类。分类方法流程图如图 2 所示。

3 实验与分析

为验证结合最小噪声分离变换和卷积神经网络分类方法的有效性, 将该分类方法与传统神经网络分类方法和支持向量机分类方法进行比较。实验数据为四川省成都市某街区地物复杂的高分一号影像, 其大小为 600 像素 \times 600 像素, 研究区域主要的地物包括道路、房屋、水体、草地、树木和阴影等。借助遥感图像处理平台 ENVI 对影像进行预处理, 并以真彩显示原始影像, 如图 3(a) 所示。在选取训练样本过程中, 综合考虑研究区的地物分布, 每类地物选取 1080 个像素点, 并全部用作有标签训练。由于影像库标签的训练数据较少, 可以很好地体现非监督预训练方法的优越性。为了对后续影像分类进行评价, 均匀选取各类地物测试样本, 结果如图 3(b) 所示。

本研究采用的卷积神经网络模型一共有 5 层, 包括 2 个卷积层、2 个池化层和 1 个全连接层。第 1 层的卷积滤波器设置为 800 个; 第 2 层卷积滤波器设置为 1080 个, 其大小为 5×5 ; 将采样间隔设置为 1, 并采用概率最大化进行采样。卷积神经网络激活函数为 $g(x) = \max(x)$ 。实验的实现平台为 Matlab, 计算机硬件配置为 2.4 GHz 主频, 4 GB 内存。分类结果如图 4(a) 所示。为验证所提分类方法的有效性, 将其与传统浅层机器学习的支持向量机分类方法和人工神经网络分类方法进行比较。支持向量机分类方法的核函数是径向基函数, 利用交叉验证的网格搜索法确定最佳的惩罚系数为 58, 不敏感系数为 0.45, 分类结果如图 4(b)

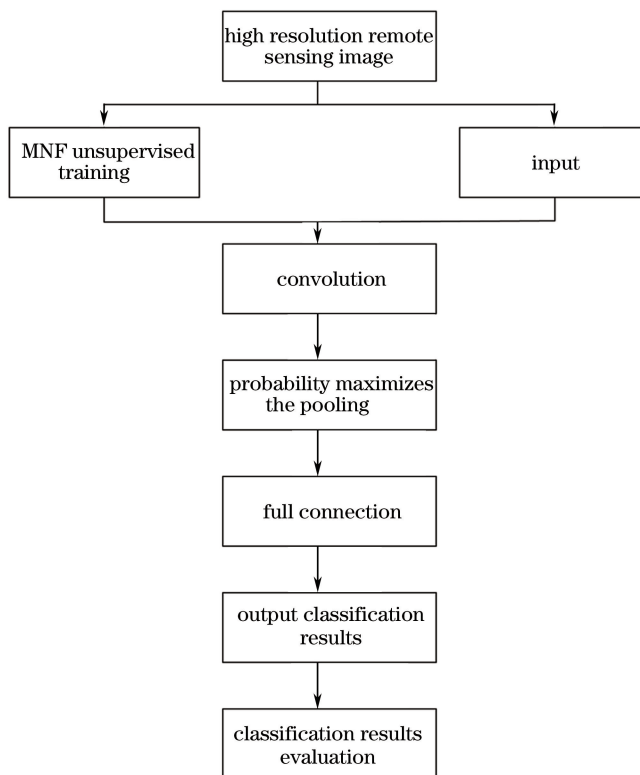


图2 结合最小噪声分离变换和卷积神经网络分类方法流程图

Fig. 2 Flow chart of proposed classification method combining minimum noise fraction rotation and convolution neural network

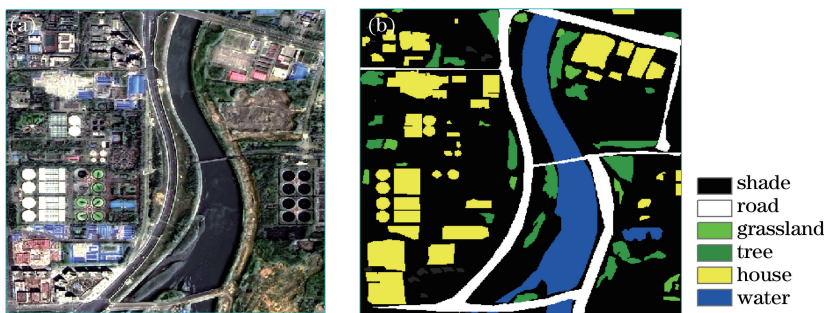


图3 (a)原始影像和(b)测试数据

Fig. 3 (a) Original image and (b) testing data

所示。人工神经网络分类结果如图4(c)所示。

由图4可知,传统的支持向量机分类方法和人工神经网络分类方法的结果均存在“胡椒粉”现象,而所提分类方法得到了清晰的地物。卷积神经网络模型因包含复杂的多层非线性变换而具有更强的表达与建模能力,因此处理高分辨影像中复杂的地物信息时能获得良好的效果。在上述三种分类方法中,所提分类方法的分类结果最优,支持向量机分类方法的分类结果最差。为了客观地定量评价所提分类方法的精度,通过计算混淆矩阵、总分类精度进行精度统计,分类精度统计结果和分类所需时间如表1所示。

由表1可知,在三种分类方法中,所提分类方法的总分类精度是最高的,说明本文所提分类方法能很好地改善高分辨影像的分类精度。对单个地物的分类精度进行分析后可知,所提分类方法的分类精度高于传统浅层机器学习分类方法,这是因为深层次的结构模型包含复杂的多层非线性变换而使其具有更强的表达与建模能力,从而更适用于处理复杂地物。结合最小噪声分离变换和卷积神经网络的高分辨率影像分类方法能获得较好的分类结果,原因在于卷积神经网络作为一种深层的网络结构,能很好地获取高分辨影像中地

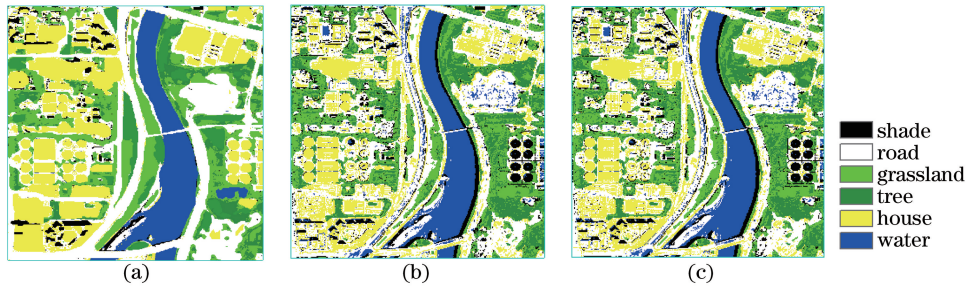


图4 不同分类方法的分类结果。(a)所提分类方法;(b)支持向量机分类方法;(c)神经网络分类方法
Fig. 4 Classification results of different methods. (a) Proposed classification method; (b) support vector machine classification method; (c) neural network classification method

表1 不同分类方法的分类精度和分类所需时间

Table 1 Classification accuracy and execution time of different methods

Classification method	Classification accuracy / %						Overall accuracy / %	Execution time / s
	Shade	Road	Grassland	Tree	House	Water		
Proposed method	96.26	95.41	95.31	95.62	94.32	96.02	95.26	120
Neural network	92.25	93.63	90.61	93.78	89.51	89.12	88.93	148
Support vector machine	85.31	93.17	85.31	78.37	85.49	90.31	87.41	108

物的本质特征,而不像浅层人工神经网络一样随机分配权值。综上所述,所提分类方法能很好地提高分类精度,并获得地物清晰的分类结果。因为卷积神经网络的激活函数为线性函数,这使得训练速度更快,所以相对人工神经网络而言,所提分类方法的时间复杂度较低。

4 结 论

针对高分辨影像,提出了结合最小噪声分离变换和卷积神经网络的分类方法,并将该方法与人工神经网络分类方法和支持向量机分类方法进行比较。卷积神经网络作为一种深层的网络结构,能很好地获取高分辨影像中的地物本质特征,采用最小噪声分离分析非监督训练滤波器,提高了卷积神经网络对地物的分类精度。所提分类方法可在较少标签数据的情况下完成卷积神经网络的初始化,但该方法仍存在许多不足,如没有判断卷积神经网络的层次结构和参数设置是否最优。下一步工作将借鉴已有的网络模型进行研究,以获得更高的分类精度。

参 考 文 献

- [1] Cheriyyadath A M. Unsupervised feature learning for aerial scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 439-451.
- [2] Luo B, Zhang L. Robust autodual morphological profiles for the classification of high-resolution satellite images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(2): 1451-1462.
- [3] Huang X, Zhang L. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(1): 257-272.
- [4] Han J, Zhang D, Cheng G, *et al.* Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(6): 3325-3337.
- [5] Liu Dawei, Han Ling, Han Xiaoyong. High spatial resolution remote sensing image classification based on deep learning[J]. Acta Optica Sinica, 2016, 36(4): 0428001.
刘大伟,韩玲,韩晓勇.基于深度学习的高分辨率遥感影像分类研究[J].光学学报,2016,36(4):0428001.
- [6] Wu Yiquan, Tao Feixiang. Multispectral and panchromatic image fusion based on improved projected gradient NMF in NSST domain[J]. Acta Optica Sinica, 2015, 35(4): 0410005.

- 吴一全, 陶飞翔. 改进投影梯度 NMF 的 NSST 域多光谱与全色图像融合[J]. 光学学报, 2015, 35(4): 0410005.
- [7] Chen Yang, Fan Rongshuang, Wang Jingxue, *et al.* Segmentation of high-resolution remote sensing image combining phase consistency with watershed transformation[J]. Laser & Optoelectronics Progress, 2017, 54(9): 092803.
陈洋, 范荣双, 王竞雪, 等. 结合相位一致和分水岭变换的高分辨率遥感影像分割方法[J]. 激光与光电子学进展, 2017, 54(9): 092803.
- [8] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]. Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012: 1097-1105.
- [9] Ranzato M A, Fergus R, Kavukcuoglu K, *et al.* Learning invariant features through topographic filter maps[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2009: 1605-1612.
- [10] Ciresan D C, Meier U, Gambardella L M, *et al.* Convolutional neural network committees for handwritten character classification[C]. Proceedings of the 2011 International Conference on Document analysis and Recognition, 2011: 1135-1139.
- [11] Kim Y. Convolutional neural networks for sentence classification[Z/OL]. (2014-09-03)[2017-00-00]. http://xueshu.baidu.com/s?wd=paperuri%3A%287ea81182039becbb82a22aaae8099c15%29&filter=sc_long_sign&tn=SE_xueshusource_2kduw22v&sc_vurl=http%3A%2F%2Ffarxiv.org%2Fabs%2F1408.5882&ie=utf-8&sc_us=17214068522800736720.
- [12] Ciresan D C, Meier U, Masci J, *et al.* Flexible, high performance convolutional neural networks for image classification[C]. Proceedings of the 22nd International Joint Conference on Artificial Intelligence, 2011: 1237-1242.
- [13] Xiao Jinsheng, Liu Enyu, Zhu Li, *et al.* Improved image super-resolution algorithm based on convolutional neural network[J]. Acta Optica Sinica, 2017, 37(3): 0318011.
肖进胜, 刘恩雨, 朱力, 等. 改进的基于卷积神经网络的图像超分辨率算法[J]. 光学学报, 2017, 37(3): 0318011.
- [14] Maggiori E, Tarabalka Y, Charpiat G, *et al.* Convolutional neural networks for large-scale remote-sensing image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(2): 645-657.
- [15] Du Y, Bourlai T, Dawson J. Automated classification of mislabeled near-infrared left and right iris images using convolutional neural networks [C]. IEEE 8th International Conference on Biometrics Theory, Applications and Systems, 2016: 16555730.