

论文引用格式:

深度卷积神经网络特征提取用于地表覆盖分类

张伟^{1, 2}, 郑柯¹, 唐婷¹, 赵理君¹

1.中国科学院遥感与数字地球研究所, 北京 100101; 2.中国科学院大学, 北京 100049

摘要: **目的:** 地表覆盖监测是生态环境变化研究、土地资源管理和可持续发展的重要基础, 在全球资源监测、全球变化检测中发挥着重要作用。提高中等分辨率遥感影像地表覆盖分类的精度具有非常重要的意义。**方法:** 近年来, 深度卷积神经网络在图像分类、目标检测和图像语义分割等领域取得了一系列突破性的进展, 相比于传统的机器学习方法具有更强的特征学习和特征表达能力。基于其优越的特性, 本文进行了深度卷积神经网络对中等分辨率遥感影像进行特征提取和分类的探索性研究。以 GF-1 的 16 米空间分辨率多光谱影像为实验数据, 利用预训练好的 AlexNet 深度卷积神经网络模型进行特征提取, 以 SVM 为分类器进行分类。分析了 AlexNet 不同层的特征以及用于提取特征的邻域窗口尺寸对分类结果的影响, 并与传统的单纯基于光谱特征和基于光谱+纹理特征的分类结果进行对比分析。**结果:** 结果表明在用 AlexNet 模型提取特征进行地表覆盖分类时, Fc6 全连接层是最有效的特征提取层, 最佳的特征提取窗口尺寸为 9×9, 同时利用深度特征得到的总体分类精度要高于其它两种方法。**结论:** 深度卷积神经网络可以提取更精细更准确的地表覆盖特征, 得到更高的地表覆盖分类精度, 为地表覆盖分类提供了参考价值。

关键词: 卷积神经网络, AlexNet, 特征提取, 地表覆盖分类, 支持向量机

Preliminary study on land cover classification with features extracted by deep convolutional neural network

Zhang Wei^{1,2}, Zheng Ke¹, Tang Ping¹, Zhao Lijun¹

1. Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing, 100101; 2. University of Chinese academy of Sciences, Beijing, 100049

Abstract: **Objective:** Monitoring land cover information is fundamental for environmental change studies, land resource management and sustainable development, which plays an important role in the global resources monitoring and change detection. Improving the land cover classification accuracy on moderate resolution remote sensing images is of great significance. For land cover classification, some researchers have combined the texture features with the spectral values. The classification accuracy has improved, but the ability of using texture features to improve the classification accuracy is still very limited. How to extract more expressive features is the key to the remote sensing image classification. In recent years, the deep convolutional neural network (CNN) has made a great breakthrough in the fields of image classification, object detection, image semantic segmentation and so on. It has a more powerful

基金项目: 高分辨率对地观测系统重大专项 (03-Y20A04-9001-15/16); 高分辨率对地观测系统重大专项 (11-Y20A05-9001-15/16)

收稿日期: ; **修回日期:**

第一作者简介: 张伟 (1992-), 男, 中国科学院遥感与数字地球研究所电子与通信工程专业硕士研究生, 主要研究方向为遥感图像分类, 机器学习与数据挖掘。E-mail: zhangw@radi.ac.cn

通信作者: 郑柯, 研究员, E-mail: zhengke@radi.ac.cn

Supported by: Major Project of High Resolution Earth Observation System of China (03-Y20A04-9001-15/16); Major Project of High Resolution Earth Observation System of China (11-Y20A05-9001-15/16)

ability of feature learning and representation compared with the traditional machine learning methods. It needs to be investigated whether the powerful ability of feature extraction of CNN can improve the accuracy of land cover classification. **Method:** Due to CNN's superior characteristics, this paper has done an exploratory research on the classification of moderate resolution remote sensing images using features extracted by deep CNN. In detail, the GF-1 multi-spectral remote sensing imagery with 16m resolution is used as experimental data, and the pre-trained AlexNet is used for feature extraction, with the support vector machine (SVM) as the classifier. The proposed method has three steps: 1). Preprocessing: because the pre-trained AlexNet was trained on ImageNet, the image that is input to AlexNet must include three bands. Thus the principal component analysis is applied to the GF-1 four-band multi-spectral image to derive the first three principal components. 2). Deep features extraction: for each pixel, the surrounding image patch with a fixed window size is firstly upsampled to 224×224 by the nearest neighbor interpolation method to conform to the required input size of the AlexNet model and then subtracts the mean image of train set for normalization, and is finally fed into the pre-trained AlexNet to extract the deep features of convolutional layers or full-connected layers. 3). Classification: the extracted deep features are fed into the SVM classifier with a linear kernel. In this study, the ability of feature representation for various layers of AlexNet and the effectiveness of different window sizes are evaluated and analyzed. Furthermore, the comparisons with classification results obtained by spectral values and spectral-texture features are conducted to assess the potential of the proposed method. **Result:** Experimental results show: 1). Features of full-connected layers are more effective than those of convolutional layers and the best feature extraction layer of AlexNet is the 6th full-connected layer. 2). The window size for feature extraction has an influence on classification results. As the size increases, the accuracy increases initially and decreases in the end. The best window size is 9×9 . 3). The classification accuracy obtained by using deep features of AlexNet is higher than those done by using only spectral values and spectral-texture features for classification. 4) A shortage of the proposed method: because of the ReLU nonlinear activation function and max pooling in AlexNet, there is a phenomenon of expansion or shrinkage of actual outline for the classes with very high spectral values or very low spectral values. **Conclusion:** In conclusion, the deep CNN can extract more accurate features for land covers and obtain a higher classification accuracy, providing a valuable reference for land cover classification. In future work, some further measures will be taken to improve the proposed method such as selecting effective activation function and pooling method to build and train a new CNN model for land cover classification.

Key words : convolutional neural network(CNN), AlexNet, feature extraction, land cover classification, SVM

0 引言

地表覆盖监测是生态环境变化研究、土地资源管理和可持续发展的重要基础^[1], 在全球资源监测、全球变化检测中发挥着重要作用。其中, 基于十米级中等分辨率遥感影像地表覆盖的计算机自动化分类技术研究至今已有非常丰富的研究结果。近年来全球尺度的地表覆盖遥感制图成为地表覆盖分类研究的热点。2013 年宫鹏等人^[2]利用 Landsat TM 和 ETM+ 数据和计算机监督分类的方法第一次获得了 30 米空间分辨率全球尺度的地表覆盖分类结果。在宫鹏等人^[2]的文章中, 分类特征主要采用光谱特征, 分类方法采用了四类方法并对分类结果进行了比较, 相比 RF(Random Forest)、J4.8 决策树、MLC(Maximum Likelihood Classifier)分类方法, SVM 分类取得了最高的分类精度, 总体分类精度达到 64.89%, 其中耕地的用户精度是 45.27%, 草地和灌木地的分类精度不足 50%, 不透水区域的分类精度也只有 30.77%, 裸地的分类精度是 62.72%。因此研究提高中等分辨率遥感影像地表覆盖分类的精度仍然具有非常重要的意义。

对中等分辨率遥感影像的地表覆盖分类, 在利用像元光谱特征分类的基础上, 学者们提出加入纹理特征进行分类^[3-4]。尽管分类精度有所提高, 但纹理信息对提高分类精度的作用仍然非常有限^[5], 分类结果很难达到满意的效果。因此如何提取出更具类别表达性的特征仍然是遥感影像分类的关键。

近年来, 卷积神经网络(Convolutional Neural Networks, CNN)在图像分类^[6]、目标检测^[7-8]、语义分割^[9]和人脸识别^[10]等领域取得了一系列突破性的研究成果。卷积神经网络将特征提取和分类结合为一体, 其局部连接、权值共享及池化操作等特性使之可以有效地减少训练参数的数目, 降低网络复杂度, 使模型对图像的平移、扭曲、缩放具有一定程度的不变性, 并具有强鲁棒性和容错能力^[11], 与传统机器学习方法相比具有更强大的特征学习和特征表达能力^[12]。卷积神经网络在高分辨率的遥感影像场景识别方面获得了成功的应用^[13-17], 已经证明了卷积神经网络强大的特征提取能力。遥感影像的场景识别与土地利用分类有很高的相似性, 都需要构建与场景语义相关的特征进行分类识别, 与地表覆盖分类有很大的不同。地表覆盖分类一般着重利用地表覆盖的光谱、纹理等中、低级特征对地物进行

分类, 面对的困难主要是类内谱的多样性引起的特征表达的困难及因混合像元引起的混合光谱的特征表达问题。

卷积神经网络强大的特征提取能力能否用于提高中分辨率遥感影像地表覆盖分类的精度? 相关研究目前非常少。针对这种情形, 本文进行了卷积神经网络对中分辨率遥感影像进行特征提取和分类的探索性研究, 以 GF-1 的 16 米空间分辨率的多光谱影像为实验数据, 用卷积神经网络 AlexNet^[18]进行特征提取, 再利用 SVM 进行分类, 并与单纯基于光谱特征的 SVM 分类结果和基于光谱+纹理特征的 SVM 分类结果进行了比较。实验结果表明卷积神经网络可以提取更精细更准确的地表覆盖特征, 得到更高的地表覆盖分类精度。实验结果也表明了 AlexNet 在地表覆盖分类中对高亮度和低亮度地物特征提取方面的不足。

1 AlexNet及其用于地表覆盖分类

1.1 AlexNet

AlexNet 是一个典型深度卷积神经网络模型。该模型在 2012 年的 ImageNet 大规模图像识别竞赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)^[19]的图像分类比赛中获得了冠军, 在图像识别领域实现了重大突破。图 1 是 AlexNet 模型的结构示意图。它是一个 8 层的卷积神经网络, 包括 5 个卷积层和 3 个全连接层。网络的最前端是输入图片的原始像素点, 最后端是图像的识别结果。

该模型在进行图像识别的过程中涉及到卷积(Conv)、池化(Pool)、规范化(Norm)运算和 Dropout 策略的使用, 下面分别简述以便了解整个网络结构。

设卷积神经网络的输入为原始影像 P , 用 F_i 表

示第 i 层的特征图, 图 1 中, $F_0 = P$ 是输入层, $F_1 - F_5$ 对应 5 个卷积层, $F_6 = Fc6$, $F_7 = Fc7$, 最后一层是输出层。假设 $F_i (1 \leq i \leq 5)$ 表示卷积层, 则 F_i 的产生过程可以用公式(1)描述为:

$$F_i = f(F_{i-1} \otimes W_i + b_i) \quad (1)$$

式中: W_i 表示第 i 层卷积核的权值向量, b_i 表示第

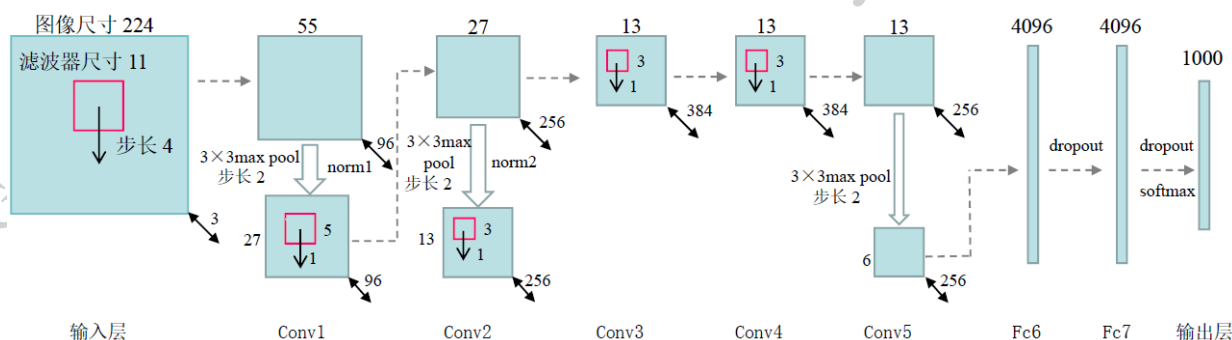


图 1 AlexNet 结构图

Fig.1 The architecture of AlexNet

i 层的偏移向量, \otimes 表示用卷积核对第 i 层特征图进行卷积操作, 最后通过非线性激励函数 f 得到第 i 层的特征图 F_i 。AlexNet 模型采用 ReLU(Rectified

Linear Units)来取代传统的 Sigmoid 和 tanh 函数作为神经元的非线性激励函数。ReLU 函数的数学表达式为公式(2):

$$f(x) = \max(0, x) \quad (2)$$

Pool 表示池化运算。池化层通常是跟随在卷积层之后的, 依据一定的下采样规则对特征图进行下采样。池化可以对特征图降维并在一定程度上保持特征的尺度不变性, 同时可以防止过拟合。假设 F_i 为池化层, 则池化的操作可以用公式(3)描述为:

$$F_i = \text{subsampling}(F_{i-1}) \quad (3)$$

式中, subsampling 为下采样函数, 包括最大值池化、均值池化或随机池化等^[20-21]。AlexNet 仅在模型的 Conv1、Conv2 与 Conv5 三个卷积层后进行最大值池化。

Norm 表示局部响应的规范化处理。AlexNet 采用的局部响应的规范化处理, 本质上是一个平滑操作, 可用公式(4)表达。

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} (a_{x,y}^j)^2 \right)^\beta \quad (4)$$

式中 $a_{x,y}^i$ 表示输入的 (x,y) 位置做第 i 次卷积并通过 ReLU 单元的结果, n 是指相同位置的第 i 次前后附近的 n 次卷积, 而 N 是总的卷积次数。 k, α, β 都是预设的一些超参数, 式中 $k=2, n=5, \alpha=10^{-4}, \beta=0.75$ 。即原来的激活 $a_{x,y}^i$ 除一个归一化权重 (分

母部分) 生成了新的激活 $b_{x,y}^i$, 相当于在同一个位置 (x,y) 不同的特征图上的激活进行了平滑。

Dropout 策略, 是对于每一个隐层的输出, 以 50% 的概率将它们设置为 0, 不再对前向或者后向的过程起任何作用。对于每一个输入来说, 使用不同的网络结构, 但权重共享, 这样求得的参数能够适应不同的情况下的网络结构, 因此提高了系统的泛化能力。AlexNet 在最后的两个全连接层中使用了该策略来减轻过拟合问题, 提高模型的泛化能力。

总结图 1 的过程, 可以说 AlexNet 的输入层经过多个卷积层和下采样层的交替传递后输入到全连接层。全连接层整合卷积层或者池化层中具有类别区分性的局部信息, 将最后一层全连接层的输出值传递给输出层, 采用 softmax 逻辑回归进行分类得到基于输入的概率分布, 从而实现图像的识别。

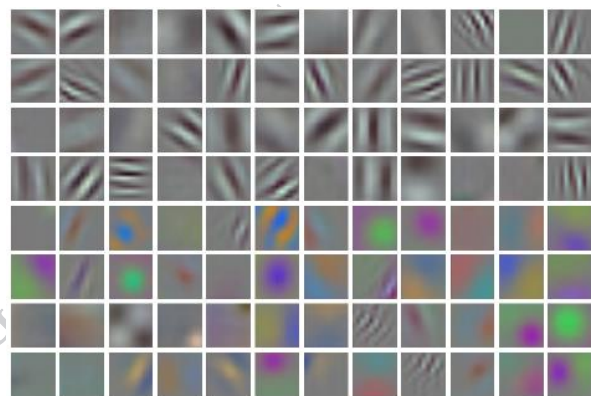


图 2 AlexNet 模型 Conv1 层中卷积核的可视化图

Fig.2 Visualization of convolution kernels in Conv1 of AlexNet

显然, AlexNet 从 F_1 - F_7 都在进行着特征提取。其中卷积层通过卷积操作和非线性激励函数获得主要特征。图 2 所示是 AlexNet 第一个卷积层中用到的 96 个 11×11 卷积核的可视化效果图^[16], 从图中可以看出第一个卷积层主要是对边缘和颜色等特征信

息进行提取。在深度卷积神经网络中，底层的卷积层提取低级特征如边缘、线条、角点，而更高层的卷积层提取更高级的能够蕴涵清晰语义信息的特征，从而提高了识别精度。

1.2 AlexNet 用于地表覆盖特征提取与 SVM 分类

图 3 为基于卷积神经网络 AlexNet 和 SVM 的地表覆盖分类流程，主要包括三个阶段：

1) 遥感影像预处理阶段：由于本文采用 ImageNet 数据集上预训练好的 AlexNet 网络模型作为特征提取器，它要求输入的影像必须为三个波段。因此为了满足模型的要求，首先对原始影像进行 PCA(Principal Component Analysis)变换来实现数据的压缩降维，取前三个主成分作为输入图像进行特征提取；

2) 深度特征提取阶段：AlexNet 要求输入经过归一化的图像，大小为 224×224 。但中分辨率遥感图像的训练样本通常都是取以样本点为中心的邻域参与训练，邻域大小尺寸远小于 224×224 ，为此需

要将待分类像元为中心的邻域插值上采样到 224×224 。这里采用最近邻插值方法以尽可能保持遥感影像的光谱信息。归一化处理是在插值后的邻域图像上减去训练样本集的均值图像^[18]，然后再输入到预训练好的 AlexNet 中提取卷积层或者全连接层的深度特征。对于影像周围邻域信息不完整的图像边缘像元，用其周围的邻域进行复制和填充以获得完整的邻域信息。本文中的特征提取是在 MatConvNet^[22]深度学习平台上完成的；

3) 分类阶段：包括训练和分类两个部分。训练部分将从训练样本提取的深度特征送入 SVM 分类器中，对 SVM 分类器进行训练，得到训练好的 SVM 分类器。分类部分将待分类像元的深度特征输入到训练好的 SVM 分类器中进行分类，将分类结果的类别标签赋给中心像元，依次遍历整个影像完成地表覆盖分类。SVM 的分类算法取自 LIBSVM 工具箱，核函数为线性核函数，参数选取采用最常用的网格选取方式寻找最优解^[23]。

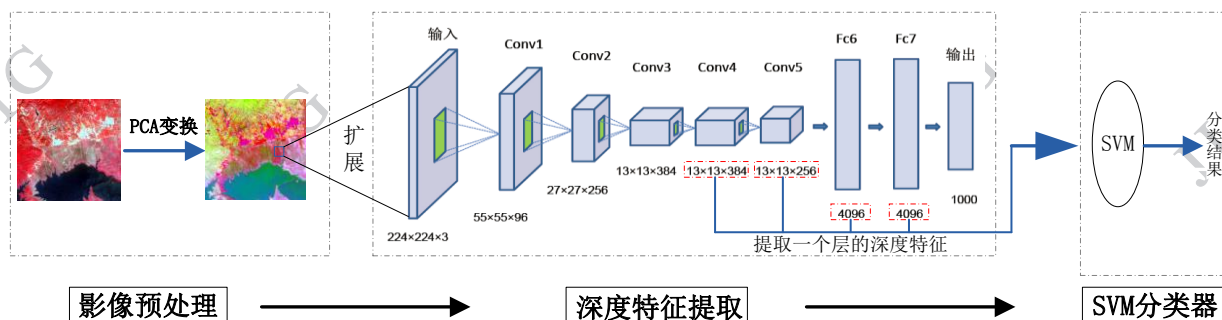


图 3 利用深度卷积神经网络特征提取进行地表覆盖分类流程图

Fig.3 The flow chart of land cover classification with deep features extracted by deep CNN

2 实验设计与结果分析

2.1 实验数据

实验数据是 GF-1 的 16 米空间分辨率多光谱影像，具体多光谱相机的技术指标如表 1 所示^[24]。图 4(a)是实验数据的假彩色合成影像，所选地区是北京市密云水库周围地区，图像大小为 500 像素×500 像素，成像时间 2016 年 6 月 22 日。根据人工解译和实地考察相结合将图中的地类分为水体、森林、居民地、耕地、休耕地、裸地 1、裸地 2 和山体阴影 8 类。根据实地考察发现裸地 1 为山上的有稀疏树木的裸露山体，而裸地 2 是通常的没有植物覆盖的裸地，其亮度在影像上较裸地 1 高。这两种地物的光谱曲线有差异，为了得到比较高的分类精度，

细分为裸地 1 和裸地 2。休耕地为耕地收割后没有种植庄稼闲置的耕地。这里的耕地类别专指其上有农作物覆盖的耕地。

表 1 GF-1 (16m) 多光谱相机技术指标

Table 1 GF-1(16m)multi-spectral camera technical parameters

波段范围 (μm)	空间分辨率 (m)	幅宽 (Km)	重访时间 (天)
0.45-0.52	16	800	2
0.52-0.59			
0.63-0.69			
0.77-0.89			

训练样本点的选择：通过人工解译和实地考察，从图像中选取了 2950 个训练样本点用于对 SVM 分类器进行训练，其中水体样本点有 400 个，森林样本点有 500 个、居民地样本点有 450 个、耕地样本

点有 400 个、裸地 1 样本点有 150 个、裸地 2 样本点有 400 个、休耕地样本点有 400 个、山体阴影样本点有 250 个。在实验的精度评价部分，同样通过人工解译和实地考察在图像中选取了 3339 个已知类别的像素点作为测试样本进行精度评价，其中水体样本点有 402 个，森林样本点有 463 个、居民地

样本点有 450 个、耕地样本点有 678 个、裸地 1 样本点有 290 个、裸地 2 样本点有 394 个、休耕地样本点有 390 个和山体阴影样本点有 272 个。验证样本的分布图如图 4(b)所示。通过计算混淆矩阵求总体分类精度和 Kappa 系数进行分类结果的精度评价。

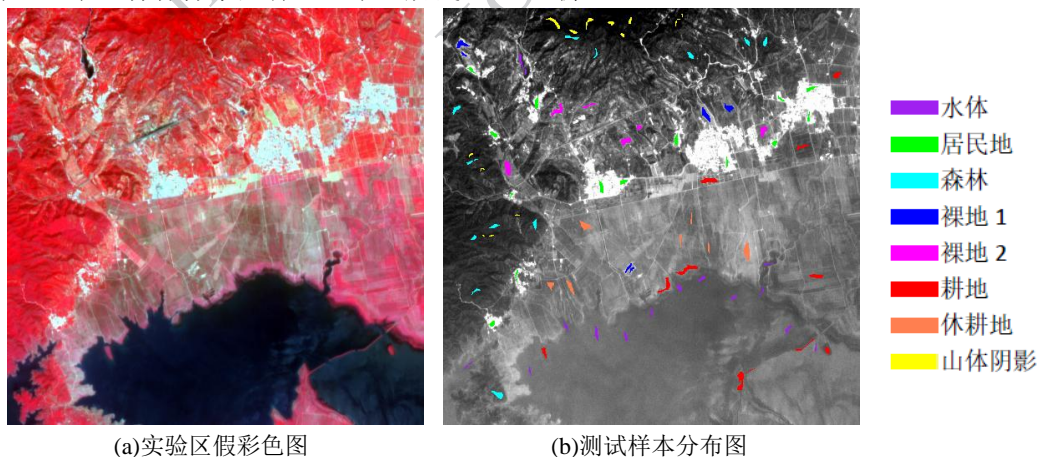


图 4 研究区域示意图

Fig.4 Images of the experimental area

((a)GF-1 false color image of the experimental area; (b)Distribution of test samples(The base map is based on blue band))

2.2 实验讨论

2.2.1 不同层特征对分类结果的影响分析

为了分析 AlexNet 模型中哪个层的特征具有更强的表达能力，从中提取最后两个全连接层 Fc6 与 Fc7 和最后两个卷积层 Conv4 与 Conv5 共四个层的特征，邻域窗口尺寸选择 9×9 ，进行分类实验并做精度评价。

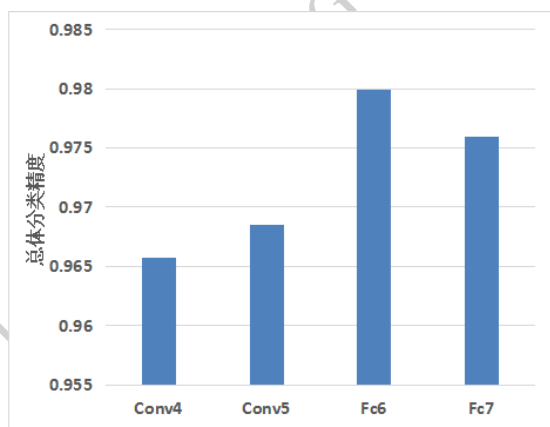


图 5 不同层特征分类结果精度对比图

Fig.5 The classification accuracy comparison chart with features extracted from different layers

从图 5 的分类结果精度对比图中可以看出：随着层数深度的增加，分类精度呈上升趋势，全连接层的特征分类精度要高于卷积层的分类精度。这是

因为深层次的特征更加抽象，更具有表达性。但 Fc6 得到的分类精度要高于 Fc7，这是因为预训练好的 AlexNet 模型是在 ImageNet 自然图像上训练得到的，Fc7 的特征虽然语义性更强，但是更多地是符合训练集的分类属性，所以分类精度反而降低。因此对于非训练集上的遥感影像，提取 Fc6 的特征反而更具表达性。

2.2.2 邻域窗口尺寸对分类结果影响分析

为了分析窗口尺寸对分类结果的影响，选出最优的窗口尺寸，提取 Fc6 的特征，选择 5×5 、 7×7 、 9×9 、 11×11 、 13×13 、 15×15 和 17×17 七个窗口尺寸对分类结果进行分析。图 6 是分类结果精度对比图。

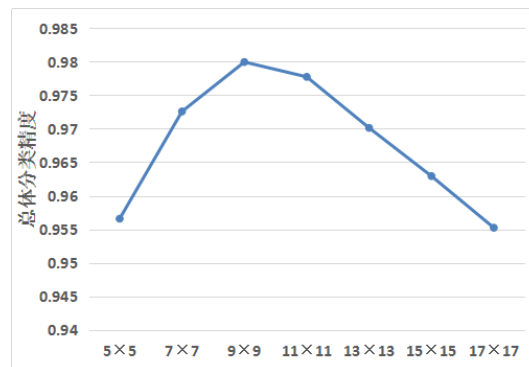


图 6 不同邻域窗口尺寸的分类结果精度图

Fig.6 The classification accuracy comparison chart with different window sizes

从图中可以看出：随着窗口尺寸的增大，精度不断提高，在 9×9 达到最大值，尺寸再增加时精度会下降。这是因为窗口尺寸较小时，邻域信息太少，不能很好提取表达地物的特征，所以精度较低；而窗口尺寸较大时，包含的信息过多，存在很多冗余信息会对分类结果造成影响，反而使分类精度降低。

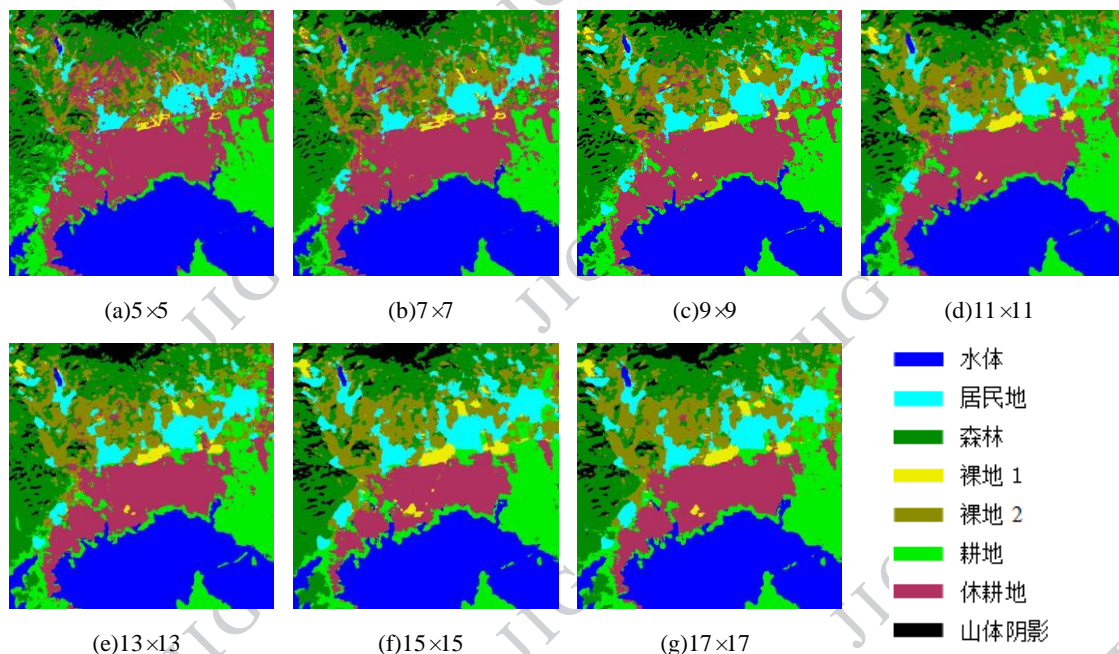


图 7 不同邻域窗口尺寸分类结果图

Fig.7 Classification images obtained with different window sizes

((a) 5×5 ; (b) 7×7 ; (c) 9×9 ; (d) 11×11 ; (e) 13×13 ; (f) 15×15 ; (g) 17×17)

2.3 实验结果对比与分析

通过 3.2 的分析可知当邻域窗口尺寸为 9×9 ，提取 Fc6 全连接层的特征时可以得到最优的分类精度。为了分析本方法的有效性，使用同一组训练样本和验证样本，进行两组方法的对比试验：

1)方法 1：采用光谱特征和 SVM 分类器分类；

2)方法 2：采用光谱特征和纹理特征，利用 SVM 分类器分类。根据纹理特征选用基于灰度共生矩阵每个波段选择两个纹理度量均值(Mean)和非相似性(Dissimilarity)，得到一个 8 维的纹理特征^[25]，然后将纹理特征和光谱特征一起输入 SVM 分类器进行分类。

从图 9 的分类结果对比图可以看出方法 1 和方法 2 的分类图中“椒盐”噪声较多，有很多细碎的小图斑，使图面杂乱，存在大量将森林错分为耕地的现象；方法 2 由于增加了纹理信息所以小图斑相对较少，分类效果有所提升；而本文方法的分类图中“椒盐”噪声较少，分类效果更好，错分漏分现象很少，能够保持分类结果中图斑的连续性，可省

从图 7 的不同尺寸分类结果图中可以看出，随着邻域窗口尺寸的增大，分类图的“椒盐”噪声的细碎图斑越来越少，分类图斑的连续性越好，分类精度增大，但是窗口尺寸过大，会造成图斑也越来越大，分类精度反而变差，这与精度评价的结果相符。

去分类的后处理操作。

在定量评价中，从图 8 中可以看出本文方法的总体分类精度和 Kappa 系数都高于都要高于其他两种方法。从图 10 的混淆矩阵的对比图中也看出，对于单个类别的分类精度，方法 1 和方法 2 都出现严重的误分现象，分类精度较低。而本文方法的精度都要高于方法 1 和方法 2，尤其对裸地 1 和裸地 2、

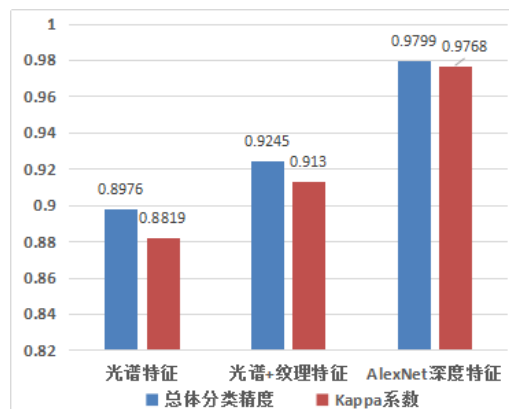


图 8 不同方法分类精度对比图

Fig.8 The accuracy comparison chart by different methods

森林和耕地这两组光谱特征差异很小的地物。究其原因估计是因为本文方法可从邻域中提取深层特征，使像元邻域光谱特征中的微小差异能够通过卷积神经网络提取的深层特征加以区分，从而使光谱特征差异很小的地物也能很好地区分。

通过定量评价得到的结论是本文方法的总体分类精度和 Kappa 系数都高于其他两种方法，方法 2 相对于方法 1 加入纹理信息使精度有所提升。但仔细观察分类结果的细节可以发现：本文方法提取的居民地范围比真实的居民地扩大了一圈，如浅蓝色

框的居民地所示；本文方法提取的小水库面积却比实际水体缩小了一圈，如黄色框所示。这可能是因为在 AlexNet 模型中应用了 $f(x)=\max(0,x)$ 的

ReLU 非线性激励函数和最大值池化函数，使得高亮地物和低亮地物(如居民地和水体)，会出现这种提取面积增加或缩水的现象。另一方面，正是因为这种现象，使得本文方法可提取出方法 1 和方法 2 不能提取出来的细小地物，如红色框所示。

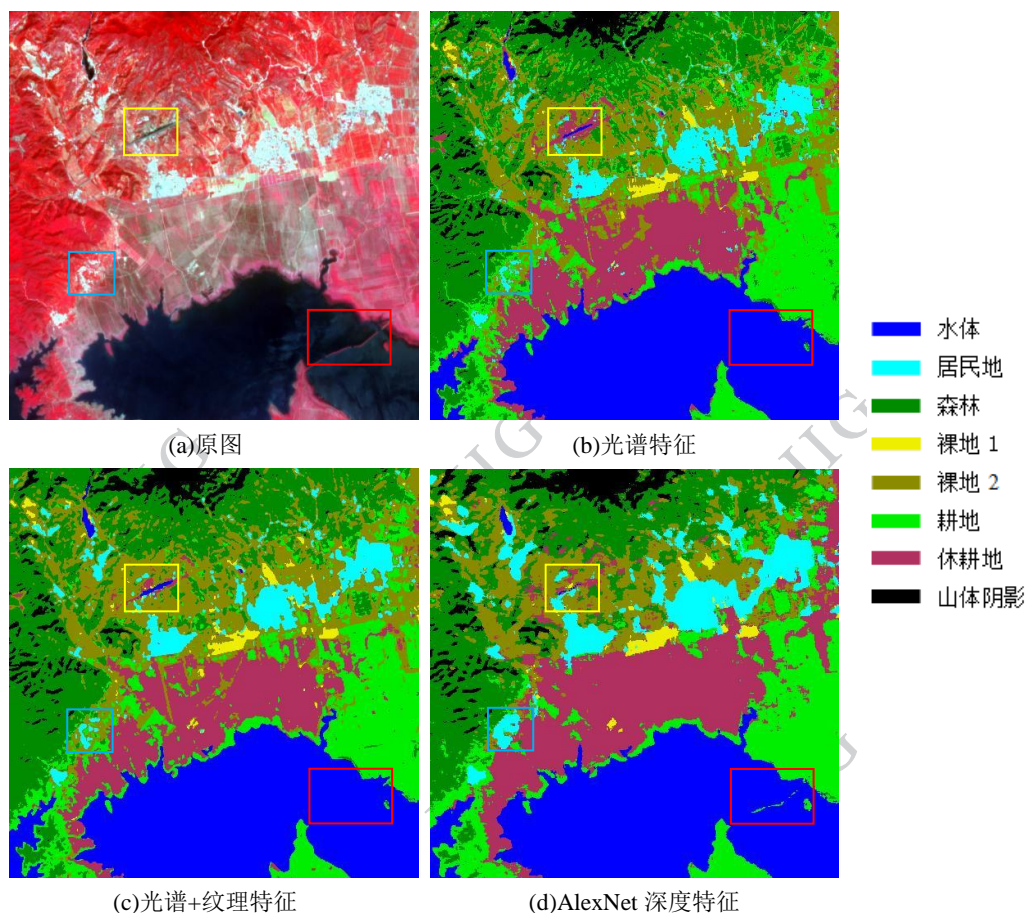
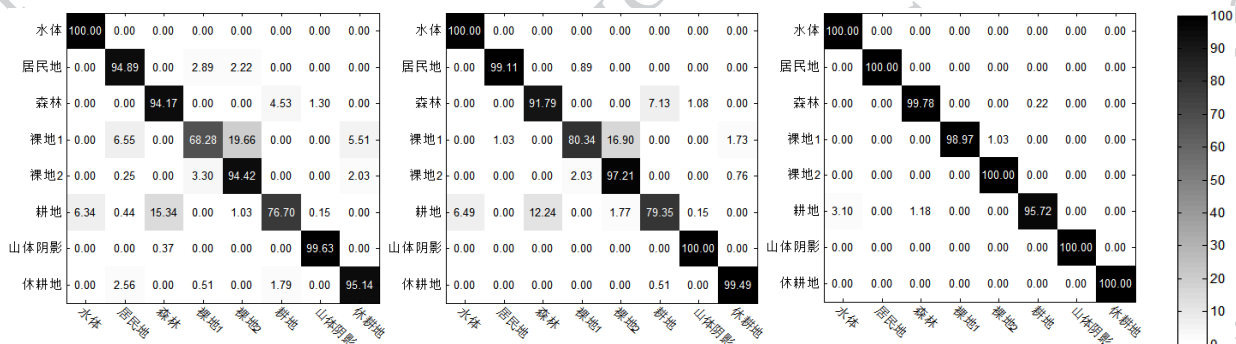


图 9 不同方法分类效果图

Fig.9 Classification images obtained by different methods

((a)Raw image; (b)Spectral values; (c)Spectral-texture features; (d)Deep features of AlexNet)



(a)光谱特征

(b)光谱+纹理特征

(c)AlexNet 深度特征

图 10 不同方法混淆矩阵对比图

Fig.10 Confusion matrices obtain by different methods

((a)Spectral values; (b)Spectral-texture features; (c)Deep features of AlexNet)

3 结论与展望

基于深度卷积神经网络强大的特征学习和特征表达能力, 本文进行了卷积神经网络对中分辨率遥感影像特征提取与分类的探索性研究。将待分类像元的邻域区域输入到预训练好的深度卷积神经网络模型 AlexNet 中提取特征, 然后将特征输入到 SVM 分类器中进行分类。在 GF-1 的 16 米空间分辨率的多光谱影像上进行了实验, 对比分析了 AlexNet 的不同层特征和邻域窗口尺寸对分类结果的影响, 并与单纯基于光谱特征和基于光谱与纹理特征的分类结果进行了对比, 得出以下结论:

1)利用深度卷积神经网络进行地表覆盖分类时, 全连接层的特征比卷积层的特征具有更强的表达性;

2)在进行基于像元的地表覆盖分类的特征提取过程中, 邻域窗口尺寸的大小对分类结果影响结果有一定影响, 随着窗口尺寸的增大, 总体分类精度先增大后减小, 因此应选择合适的窗口尺寸进行特征提取;

3)相比较于光谱特征与纹理特征, 深度卷积神经网络的深度特征更具表达性, 可以得到更高的地表覆盖分类精度, 且分类结果图中能够保持好的连续性, 无需后续处理即可取得很好的分类效果;

4)本文方法也存在不足: 由于 AlexNet 模型中应用 ReLU 作为非线性激励函数, 并使用最大值池化, 对于高亮度或低亮度的地物会出现误分和漏分的现象, 导致地物的类别边缘相对实际地物有增大或缩水的现象。

在今后的工作中, 一方面继续探索比较不同的深度卷积神经网络模型如 VGG^[26]、GoogleNet^[27]与 ResNet^[6]等用于中等分辨率遥感图像地表覆盖分类的能力; 另一方面, 将改造 AlexNet 模型, 通过选择合适的非线性激励函数、池化操作和特征层数构建和训练适合于中等分辨率遥感图像地表覆盖分类的卷积神经网络结构, 为提高中等分辨率遥感图像地表覆盖分类精度提供新的技术途径。

参考文献(References)

- [1] Chen J, Chen J, Liao A P, et al. Global land cover mapping at 30m resolution: a pok-based operational approach[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 103: 7-27. [DOI: 10.1016/j.isprsjprs.2014.09.002]
- [2]Gong P, Wang J, Yu L, et al. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data[J]. International Journal of Remote Sensing, 2013, 34(7): 2607-2654. [DOI: 10.1080/01431161.2012.748992]
- [3]Chen B, Zhang Y J, Chen L. RS image classification based on SVM method with texture[J]. Engineering of Surveying and Mapping, 2007, 16(5): 23-27.[陈波, 张友静, 陈亮. 结合纹理的 SVM 遥感影像分类研究[J]. 测绘工程, 2007, 16(5): 23-27.] [DOI: 10.3969/j.issn.1006-7949.2007.05.007]
- [4]Chen G F, Zeng G W, Chen H, et al. Study of RS image classification method based on texture features and neural network algorithm[J]. Journal of Chinese Agricultural Mechanization, 2014, 35(1): 270-274. [陈桂芬, 曾广伟, 陈航, 等. 基于纹理特征和神经网络算法的遥感影像分类方法研究[J]. 中国农机化学报, 2014, 35(1): 270-274.] [DOI: 10.3969/j.issn.2095-5553.2014.01.063]
- [5]Li Y, Li Y H, Wang J X, et al. A comparative study of SVM and ANN in multispectral image classification[J]. Hydrographic Surveying and Charting, 2016, 36(5): 19-22. [李颖, 李耀辉, 王金鑫, 等. SVM 和 ANN 在多光谱遥感影像分类中的比较研究[J]. 海洋测绘, 2016, 36(5): 19-22.] [DOI: 10.3969/j.issn.1671-3044.2016.05.005]
- [6]He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, Nevada, USA: IEEE, 2016: 770-778. [DOI: 10.1109/CVPR.2016.90]
- [7]Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus, OH, USA:IEEE,2014: 580-587. [DOI: 10.1109/CVPR.2014.81]
- [8]Girshick R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1440-1448. [DOI: 10.1109/ICCV.2015.169]

- [9]Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, Massachusetts, USA: IEEE, 2015: 3431-3440. [DOI: 10.1109/CVPR.2015.7298965]
- [10]Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[C]//Proceedings 2015 British Machine Vision Conference(BMVC). Swansea, UK: BMVA, 2015: 4101-4112. [DOI: 10.5244/C.29.41]
- [11]Zhou F Y, Jin L P, Dong J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(7): 1-23. [周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(7): 1-23.]
- [12]Lu H T, Zhang Q C. Application of deep convolutional neural network in computer vision[J]. Journal of Data Acquisition and Processing, 2016, 31(1):1-17. [卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理, 2016, 31(1): 1-17.] [DOI: 10.16337/j.1004-9037.2016.01.001]
- [13]Hu F, Xia G S, Hu J W, et al. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery[J]. Remote Sensing, 2015, 7(11): 14680-14707. [DOI: 10.3390/rs71114680]
- [14]Zhong Y F, Fei F, Zhang L P. Large patch convolutional neural networks for the scene classification of high spatial resolution imagery[J]. Journal of Applied Remote Sensing, 2016, 10(2): 025006(1-20). [DOI: 10.1117/1.jrs.10.025006]
- [15]Marmanis D, Datcu M, Esch T, et al. Deep learning earth observation classification using ImageNet pretrained networks[J]. IEEE Geoscience and Remote Sensing Letter, 2016, 13(1): 105-109. [DOI: 10.1109/lgrs.2015.2499239]
- [16]Castelluccio M, Poggi G, Sansone C, et al. Land use classification in remote sensing images by convolutional neural networks[J]. Acta Ecologica Sinica, 2015, 28(2): 627-635.
- [17]Penatti O, Nogueira K, Santos J. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Boston, Massachusetts, USA: IEEE, 2015: 44-51. [DOI: 10.1109/cvprw.2015.7301382]
- [18]Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of Advances in Neural Information Processing Systems. Lake Tahoe, USA: NIPS, 2012: 1106-1114.
- [19]Russakovsky A, Deng J, Su H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252. [DOI: 10.1007/s11263-015-0816-y]
- [20]Y-Lan B, Nicolas L R, Francis B, et al. Ask the locals: multi-way local pooling for image recognition[C]//Proceedings of the 2011 International Conference on Computer Vision (ICCV). Barcelona, Spain: IEEE, 2011: 2651-2658. [DOI:10.1109/iccv.2011.6126555]
- [21]Matthew D Z, Rob F. Stochastic pooling for regularization of deep convolutional neural networks[C]//Proceedings of 2013 International Conference on Learning Representations(ICLR). Scottsdale, Arizona, USA: ICLR, 2013: 1-9.
- [22]Vedaldi A, Lenc K. MatConvNet--Convolutional Neural Networks for MATLAB[C]//Proceeding of the 23rd ACM International Conference on Multimedia. Brisbane, Australia: ACM, 2015: 689-692. [DOI: 10.1145/2733373]
- [23]Hsu C W, Chang C C, Lin C J. A practical guide to support vector classification[R]. Taiwan: Department of Computer Science and Information Engineering, National Taiwan University, 2003.
- [24]中国资源卫星应用中心. 高分一号 [EB/OL]. 2014-10-15[2017-04-20].<http://www.cresda.com/CN/Satellite/3076.shtml>.
- [25]Lu M F, Feng W F, Xu Y, et al. Research on selection of features, bands and windows in classification based on texture feature[J]. Image Technology, 2011, (3): 54-59. [卢茂芬, 冯伍法, 胥亚, 等. 基于纹理的影像分类的特征波段与窗口选择问题研究 [J]. 影像技术, 2011, (3): 54-59.] [DOI:10.3969/j.issn.0270-2011.03.11]
- [26]Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of 2015 International Conference on Learning Representations(ICLR). San Diego, USA: ICLR, 2015: 1-14.
- [27]Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston, Massachusetts, USA: IEEE, 2015: 1-9. [DOI:10.1109/CVPR.2015.7298594]