

## 顾及局部特性的 CNN 在遥感影像分类的应用

杨嘉树, 梅天灿, 仲思东

YANG Jiashu, MEI Tiancan, ZHONG Sidong

武汉大学 电子信息学院, 武汉 430072

School of Electronic Information, Wuhan University, Wuhan 430072, China

**YANG Jiashu, MEI Tiancan, ZHONG Sidong. Application of Convolution Neural Netowrk using region information to remote sensing image classification. Computer Engineering and Applications, 2018, 54(7): 188-195.**

**Abstract:** Deep learning shows excellent performance in image recognition and natural language processing. This paper presents a learning method which is based on Convolution Neural Network (CNN) to classify remote sensing images. Aiming at the problem of low efficiency caused by too many sampling windows in remote sensing image classification, this paper proposes a novel method based on region to improve the classification efficiency. There are two steps in the classification process. The first step of the proposed algorithm is to classify the input images by convolution neural networks. Then, SVM is used to classify the results of convolution neural networks. The proposed algorithm is evaluated by being compared with the results of other existing methods. Experimental results show that proposed algorithm can achieve higher classification accuracy.

**Key words:** Convolution Neural Networks(CNN); image segmentation; image classification; transfer learning

**摘 要:**深度学习的方法在图像识别和自然语言处理等方面展示了优异的性能。将卷积神经网络(Convolution Neural Network, CNN)用于高分辨率遥感影像分类。针对CNN用于遥感影像分类使用固定大小窗口遍历时,影像采样窗口数量过多,导致的分类效率低下问题,提出一种基于影像区域特性的采样窗口确定方法,提高分类效率。影像分类包括两个阶段:首先,利用卷积神经网络得到的特征对影像进行分类;然后,采用支撑向量机对第一步分类由于特征区分性不足造成的错分地物类别进行再分类。采用具有不同特性的遥感影像对所提方法进行了验证,实验结果表明,同现有的特征表示和分类方法相比,该方法的性能有明显改善。

**关键词:**卷积神经网络(CNN);影像分割;影像分类;迁移学习

**文献标志码:**A **中图分类号:**TP753 **doi:**10.3778/j.issn.1002-8331.1611-0225

### 1 引言

遥感是一种对地实施观测和资源管理的有效手段。遥感影像分类,在土地规划、灾害救灾以及变换检测等应用中发挥着关键作用。由于高分辨率影像中普遍存在的“异物同谱”和“同物异谱”现象<sup>[1]</sup>,现有高分辨率影像分类技术不能有效满足实际应用需求,高分辨率遥感影像分类方法研究仍然是一个开放性的问题。

影像分类性能在很大程度上取决于影像数据特性的表示方式。结合影像光谱和空间信息的特征表示方法在高分影像分类中得到广泛应用,如LEWA (Length-

width Extraction Algorithm)<sup>[2]</sup>、PSI(Pixel Shape Index)<sup>[3]</sup>等算子通过描述影像的形状特征,改善了分类性能。Yuan等<sup>[4]</sup>利用局部光谱特征统计图、SIFT(Scale-Invariant Feature Transform)<sup>[5]</sup>、HOG(Histogram of Oriented Gradient)<sup>[6]</sup>算子等描述影像的空间信息用于影像分类。为进一步提高特征的鲁棒性,人们通过对底层特征进行编组处理得到中层特征(Mid-level feature)。BoVW(Bag of Visual Words)<sup>[7]</sup>是一种广泛应用的中层特征描述子:首先使用SIFT等算法提取影像特征的描述算子,生成视觉“词典”,然后统计视觉“单词”的数

**基金项目:**国家测绘地理信息公益性行业科研专项项目(No.201412015)。

**作者简介:**杨嘉树(1992—),男,硕士,主要研究领域为图像处理;梅天灿(1969—),男,博士,副教授,主要研究领域为图像处理与模式识别, E-mail: mtc@whu.edu.cn;仲思东(1963—),男,博士,教授,主要研究领域为机器视觉。

**收稿日期:**2016-11-14 **修回日期:**2017-01-11 **文章编号:**1002-8331(2018)07-0188-08

**CNKI网络优先出版:**2017-04-01, <http://kns.cnki.net/kcms/detail/11.2127.TP.20170401.0853.032.html>

目得到特征,结合boosting等方法进行分类。从满足实际应用需求的角度看,这些特征的适用范围以及描述高层语义特征的能力都有待提高。近年来,深度学习的方法在图像识别和语言处理等方面展示了优异的性能<sup>[8]</sup>。卷积神经网络(Convolution Neural Network, CNN)是一种常见的深度学习结构。卷积神经网络模拟人脑视觉分层机理,通过学习的方法得到数据特征,省去了设计特征的困难。此外,卷积神经网络提取的特征,相对于传统人工设计特征,有平移和旋转不变性,在图像分类和目标识别等领域的表现更优。卷积神经网络正在被应用于计算机视觉的众多领域。

卷积神经网络是一种数据驱动模型,需要大量的数据作为支撑。由于遥感影像数据特性的限制,直接采用现有卷积神经网络模型表示遥感影像数据特性存在以下问题:(1)基于卷积神经网络的特征学习方法需要大量训练样本。如果训练样本不足,学习得到的特征不能有效描述数据特性。遥感影像数据分类一般难于获取足够的训练样本<sup>[9]</sup>。(2)现有基于卷积神经网络的影像分类方法需要首先确定目标在影像中可能出现的位置,然后在该位置开一个窗口,最终根据窗口内的数据确定目标的类别。遥感影像分类需确定每个像素的类别,因此要以每个像素为中心开窗获取的数据确定像素的类别。这种处理使得计算效率很低,窗口大小选择不合适也会显著降低分类性能。因此,直接采用现有卷积神经网络模型进行遥感影像分类不能有效提高分类性能。

针对直接使用卷积神经网络进行遥感影像分类存在的问题,本文的主要工作:(1)利用影像分割获取的区域信息确定用于卷积神经网络的子窗口影像,提高了卷积神经网络对遥感影像分类的效率;(2)采用支撑向量机(Support Vector Machine, SVM)在卷积神经网络分类结果的基础上进行再分类,进一步提高分类精度。

2 影像分类

本文提出的遥感影像分类方法流程如图1所示。本文实现高分辨率遥感影像分类时,首先对影像使用Mean Shift<sup>[10]</sup>进行过分割,根据分割结果确定输入卷积

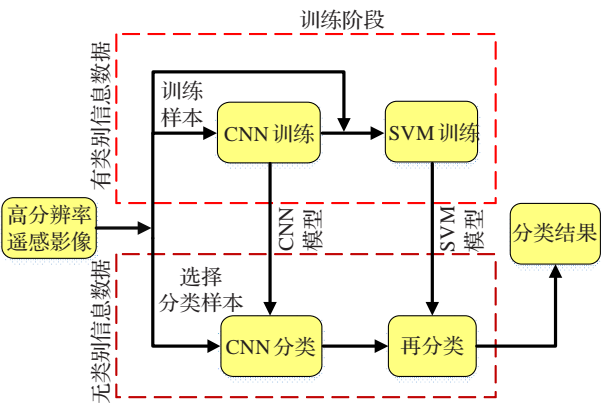


图1 分类框架流程图

神经网络的影像块,使用卷积神经网络对影像块进行分类,分类结果作为该区域的分类结果,在所有区域上使用同样操作得到初始分类结果。在此基础上,针对卷积神经网络分类的结果,使用支撑向量机进一步改善分类性能。

2.1 卷积神经网络分类框架

卷积神经网络应用于特征学习和分类的框架如图2所示。卷积神经网络包含多个用于提取特征的卷积层(convolution layer)、池化层(pooling layer)和全连接层(fully connected layer)。卷积层对输入图像或特征图进行卷积,其输出表示为 $(h_k)_{ij}=f((W_k*x)_{ij}+b_k)$ ,其中 $k=1,2,\cdots,K$ 为卷积核的索引,也是输出特征图的索引, $x$ 为第 $k$ 个卷积核的输入, $W_k$ 和 $b_k$ 分别为第 $k$ 个卷积核的权重和偏置值, $(h_k)_{ij}$ 为第 $k$ 个特征图在 $(i,j)$ 位置的响应值,“ $*$ ”表示二维卷积运算, $f(\cdot)$ 为卷积层的激活函数。卷积层的输出特征图作为卷积核在输入上的响应。池化层通常对卷积层输出中的子区域使用最大池化或平均池化进行下采样。池化操作一方面可增加特征的鲁棒性,另一方面减少了卷积神经网络的参数。全连接层使用 $y_k=\sum W_{kl}x_l+b_k$ 计算输出,其中 $x_l$ 为输入, $y_k$ 为全连接层的第 $k$ 个输出, $W_{kl}$ 和 $b_k$ 为全连接层的权重和偏置值。全连接层的存在使得卷积神经网络的输入尺寸固定。Softmax层可视为全连接层的一种,输出为数据类别概率的归一化分布。对于多分类而言,常取归一化概率最大的标签作为分类结果。

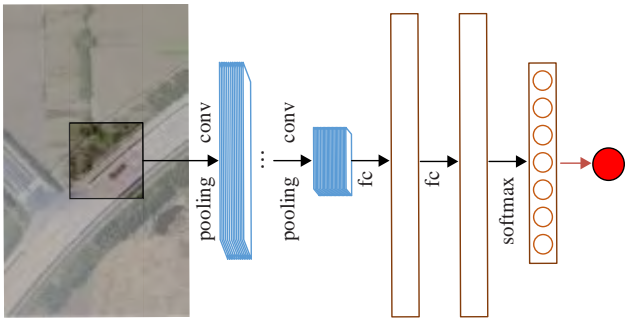


图2 卷积神经网络结构示意图

卷积神经网络中存在全连接层,全连接层节点的固定导致卷积神经网络的输入为固定尺寸。卷积神经网络应用于图像识别时,需将输入图像缩放为固定大小,如AlexNet<sup>[11]</sup>输入大小为 $224\times 224$ 。遥感影像分类需要确定影像中每个像素的类别。如图2所示,卷积神经网络用于遥感影像分类时,也需以当前像素点为中心开子窗口,将该子窗口影像数据缩放至卷积神经网络的输入尺寸后,输入网络,从而确定当前像素的类别。这种基于分块影像分类的方法是平移不变的,即像素的类别与其周围的像素有关,与其所处的位置无关。对于宽1 300像素,高1 300像素的图像,当窗口滑动步长为3,以长宽均为64像素取影像块,影像块数目超过 $10^5$ 。以



单块影像通过卷积神经网络的时间为15 ms计,用CNN对影像分类总时间超过25 min,直接应用该方法时间消耗巨大。

## 2.2 基于区域信息的卷积神经网络分类窗口数据采集方法

针对影像采样窗口过多导致效率低这一问题,本文提出一种基于影像数据局部特性的窗口数据采集方法。该方法可以显著减少采样窗口的数量,从而提高分类效率。

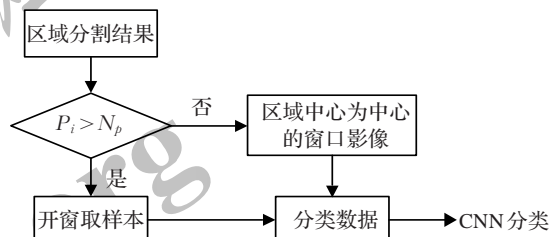
本文后续进行描述时,将影像分割后的像素集合称为区域,将放缩到卷积神经网络的输入尺寸的影像称为窗口采样影像。

影像分割是指将影像数据中有不同意义的区域或部分等分割开来,这些分割出来的区域彼此互不重叠且满足一定的区域一致性。影像分类是将影像数据按照其反映的特征分成预定义的类别并为每个类别赋予类别标记方法。影像分割后的区域在光谱信息上有一定的相似性,同时也有纹理、大小或上下文之间的一致性。通常假设每个区域中的全部像素都是属于一个类别的,实际上这种假设很难成立,因此在实际应用中常采用过分割方式确保区域内的像素具有相同的类别。由于过分割区域内像素具有相同类别,在利用卷积神经网络分类给出区域内一个像素的类别,就可得到属于该区域的所有像素类别。

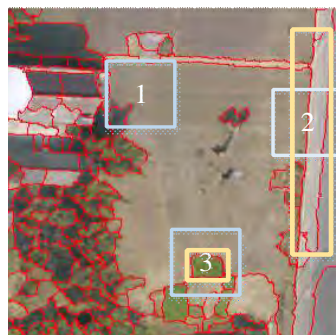
Mean Shift作为一种统计聚类方法,在图像分割中得到广泛应用。Mean Shift分割得到的区域使得区域内像素具有相同的属性,从而使得区域内像素具有相同类别。直接将Mean Shift过分割后的区域外接矩形的影像数据作为卷积神经网络的输入存在两个问题:一方面,遥感影像中不同地物大小各异导致Mean Shift分割后的区域大小不同,直接将区域缩放至同样大小会损失影像反映包含的信息,造成错分;另一方面,Mean Shift分割的区域可能为任意形状,缩放后进行分类可能引入过多其他地物类别信息,造成错分,如分割后出现环形区域等。这些问题的出现会降低分类精度。为克服以上问题,当过分割后区域面积较大,即分割的区域总像素数大于某一阈值时,选取区域中固定大小的影像块作为窗口采样影像输入卷积神经网络进行分类。当区域面积过小时,以当前区域形心为中心选取固定大小的影像块作为采样数据输入卷积神经网络进行分类。当区域面积过小时,通过取一个固定窗口数据可以保证输入卷积神经网络的数据包含足够的信息用于分类。本文采用下述策略确定采样窗口。

记Mean Shift分割后其中的第 $i$ 区域像素数目为 $P_i$ ,其外接矩形总像素数为 $R_i$ 。第 $i$ 个区域的窗口采样影像大小为 $M \times M$ (像素 $\times$ 像素)的窗口影像数据,

记为 $C_i$ ;每个区域至少包含的像素数为 $N_p$ 。 $C_i$ 中第 $i$ 区域的像素在窗口总像素比为 $w$ ,设定最小比为 $w=0.5$ 。和图3为分类样本选择过程的示例。



(a) 分类时卷积神经网络选择样本流程



(b) 分类时卷积神经网络选择样本示例

图3 分类时卷积神经网络选择样本流程和示例

(1) 若 $P_i \leq N_p$ ,则以区域中心像素为中心的 $M \times M$ 窗口影像作为卷积神经网络输入,输出作为该区域的分类结果。

(2) 若 $P_i > N_p$ ,取 $M \times M$ 窗口并以外接矩形左上角的坐标为窗口左上点坐标。若 $w > w_0$ ,取得该窗口影像为窗口采样影像;否则,向下或向右滑动 $M \times M$ 窗口,直至 $w > w_0$ 。若遍历区域后窗口采样影像中均不满足 $w > w_0$ ,则取 $w$ 最大的 $M \times M$ 窗口的影像数据作为窗口采样影像,输入卷积神经网络。将输出作为区域的分类结果。

基于Mean Shift分割区域获取卷积神经网络分类窗口的过程如图3所示。分割后采样窗口红色线为Mean Shift分割结果,淡蓝色框为每个区域对应的采样窗口,黄色矩形框对应区域的外接矩形。区域1面积较大,包含像素多, $w > w_0$ ,采样影像窗口左上角坐标与区域左上角坐标重合。区域2中 $P_i > N_p$ ,遍历区域内所有像素后 $w < w_0$ ,因此采样窗口选择在 $w$ 最大的影像块处。区域3中 $P_i < N_p$ ,以区域3形心为中心固定大小的窗口作为采样窗口。将窗口影像缩放至卷积神经网络的输入尺寸进行分类,得到分类结果。

## 2.3 基于SVM的后续分类

高分遥感影像不同地物的高度相似性使得基于卷积神经网络的分类结果仍然存在部分错分现象。通过对错分现象的分析,发现光谱信息差别很大的地物被卷积神经网络误分为一类,如道路和屋顶的光谱信息差别

较大,卷积神经网络却不能区分这两类物体。这种现象出现的原因可能是光谱信息差别较大的两类物体经卷积神经网络逐层抽象后,在新的特征空间没有可区分性。为克服这一问题,基于错分物体光谱特性差别较大这一事实,本文采用支撑向量机直接利用光谱信息对卷积神经网络易于错分的地物类别进行进一步分类。支撑向量机<sup>[12]</sup>是一种有坚实理论基础的小样本学习方法。它采用结构风险最小化准则,在最小化样本点误差的同时,缩小模型预测误差的上界,提高了模型的泛化能力。因此本文在卷积神经网络分类的基础上,使用支持向量机结合光谱特征进行再分类,以进一步提高分类精度。

3 实验与分析

3.1 实验数据与卷积神经网络训练

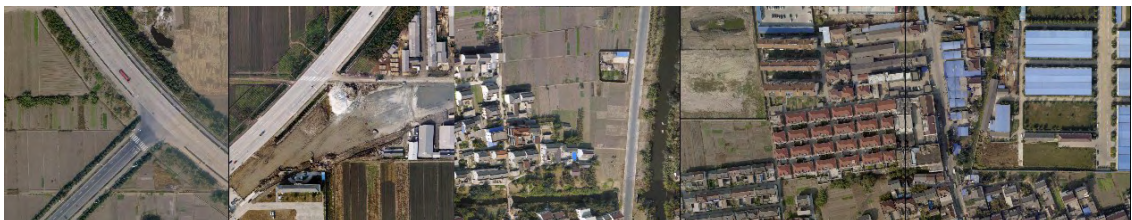
本文选取的影像数据覆盖部分江苏泰州区域,共5幅影像。每幅影像为1 300×1 300像素,RGB三通道图像,分辨率0.2 m,如图4(a)所示;图4(b)为对应影像的类别图;实验中卷积神经网络分类类别为农田、草地、树木、水体、荒地、道路、灰色屋顶、珊瑚色屋顶、蓝色屋顶9种,注释如图4(c)。

为验证方法有效性,本文设计了3个实验:(1)不同影像分割方法的卷积神经网络类结果;(2)本文层次分类结果;(3)本文与常用方法对比。目前研究者一般从视觉效果和定量指标来评价影像分类质量。从视觉效果来说,好的分类结果能保持边缘特性且具有好的区域一致性;定量指标常用混淆矩阵、分类精确度等。本文3组实验使用混淆矩阵作为精度评定的定量指标。混淆矩阵中,用户精度表示被正确标记为A类的像素个数与分类器将整个影像的像素点分为A类的像素个数的

总和的比值。生产者精度表示被正确标记为A类的像素个数与A类真实参考数目的比值。生产者精度和用户精度分别对应漏分误差和错分误差。

卷积神经网络中训练参数,需要大量数据参与训练。如本文使用的AlexNet<sup>[11]</sup>包含60多万参数,原始训练集有超过10<sup>7</sup>张图像。遥感影像分类时收集如此大量的数据不现实。训练数据不足导致卷积神经网络模型无法生成数据的准确描述。迁移学习<sup>[13]</sup>(transfer learning)旨在将对原始数据集的理解迁移到目标数据集。尤其当目标数据集数据量不足时,将源数据集上训练的分类器修改后应用到目标数据集中,可有效提高模型描述目标数据特性的能力。

ImageNet<sup>[14]</sup>中图像主体明确,如tabby cat、tiger cat等。单片遥感影像通常覆盖范围广,地物信息丰富,仅对单片影像给出一个类别意义不大,本文处理遥感影像数据时先对影像分块。分块后按照一定策略确定分块影像类别:(1)以一定像素大小的窗口在遥感影像与其标签图(ground truth)上滑动,得到分块影像集S<sub>1</sub>和对应类别集合L<sub>1</sub>。(2)设类别标签为1,2,⋯,K共K种类别,统计每小块影像中每类的像素个数n<sub>k</sub>,记采样窗口总像素为N。若对于k∈{1,2,⋯,K},存在n<sub>k</sub>>0.5N,则该小块影像的类别标签为k,若不存在符合条件的n<sub>k</sub>,则将该小块影像舍弃,得到数据集D<sub>1</sub>。(3)为增强模型的旋转不变性,再以2M×2M的窗口在影像和对应ground truth上滑动,得到分块影像集S'和对应类别集合L'。将S'旋转θ∈{30,60,90,120,150,180}后,取中心M×M窗口影像,得到小块影像集S<sub>2</sub>和对应类别集合L<sub>2</sub>。在S<sub>2</sub>中同样按步骤2得到数据集D<sub>2</sub>。将D=D<sub>1</sub>∪D<sub>2</sub>,L=L<sub>1</sub>∪L<sub>2</sub>作为训练时的样本集和对应标签。



(a)本文训练时所用影像数据示例



(b)图a影像 ground truth

■ 田地 ■ 树木 ■ 草地 ■ 荒地 ■ 道路 ■ 水体 ■ 灰色屋顶 ■ 蓝色屋顶 ■ 珊瑚色屋顶

(c)地物类别对应影像注释

图4 实验影像数据及训练、参考样本



基于卷积网络的迁移特性,本文以 AlexNet 模型为基础结构,将 softmax 层输出由 1 000 修改为 9。预训练模型为 Evan Shelhamer 使用 caffe<sup>[15]</sup>,在 ImageNet<sup>[14]</sup>数据集上训练得到的模型(top-1 准确率 57.3%, top-5 准确率 80.2%)。使用随机梯度下降方法进行 fine-tuning,反映本文数据特性。在确定训练样本窗口时,若窗口过大,丢弃的窗口数过多导致训练样本不足。若窗口过小,卷积神经网络进行训练时无法有效学习不同地物特征情况。综上所述,本文通过交叉验证的方式,选取  $M=64$  获取窗口数据。获得窗口影像后将其放大至 AlexNet 输入大小即  $224 \times 224$  进行训练。总训练样本数目约为 48 000 个,测试样本约为 5 000 个,测试精度为 96.1%。

训练样本采集流程如图 5 所示。

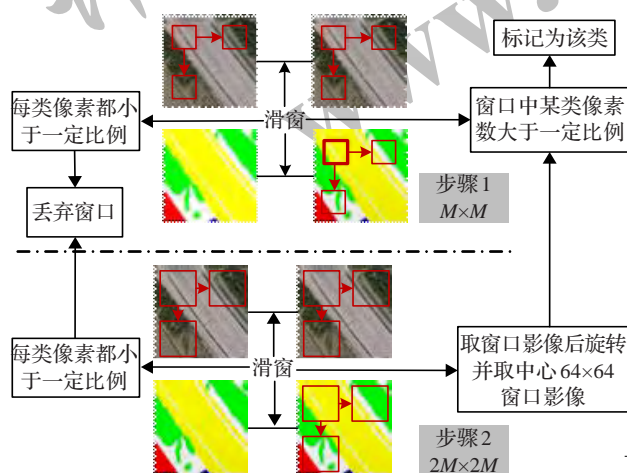


图5 训练样本采集流程

### 3.2 实验结果与分析

#### 实验1 不同采样窗口影像的分类结果对比

实验1 对比在训练影像上滑动取窗口影像, SLIC (Simple Linear Iterative Clustering)<sup>[16]</sup> 分割超像素分

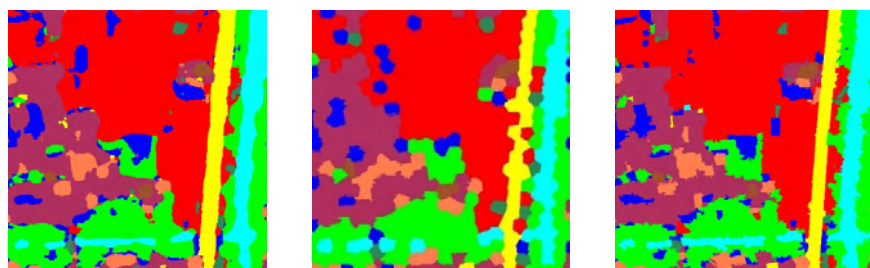
类, Mean Shift 分割后分类选取基元在卷积神经网络下的结果。第一种方法中窗口大小  $64 \times 64$ , 步长  $3 \times 3$ , 将窗口影像放大至  $224 \times 224$  后输入卷积神经网络, 分类结果作为  $3 \times 3$  小窗结果。第二种方法将超像素块初始大小设为  $64 \times 64$ , 超像素聚类结束后, 取每个超像素的外接矩形的影像得到窗口影像。得到窗口影像后, 将窗口影像放缩至  $224 \times 224$  后输入卷积神经网络, 分类结果作为该超像素区域的分类结果。实验结果如图 6 所示。

从图 6 的结果可以看出, 第一种方法边缘模糊。Mean Shift 结果边缘清晰, 目视效果最好。表 1 给出了这两个方法的精度对比。因在真实类别图中未标记蓝色屋顶和荒地类别, 故这两列数据为 0。从目视效果上看, SLIC 方法分类结果粗糙度很大, 影像的边缘信息基本消失, 效果最差, 因而没有给出混淆矩阵。从表 1 中发现, 第一种方法与第三种总体精度接近, 相差不大。从时间上看, 对于本文影像而言, 滑窗方法计算超过  $10^6$  个窗口, 远大于 Mean Shift 需要计算的窗口数 (约  $10^3$  个)。本文选用 Mean Shift 过分割得到区域信息比较合理。

在测试影像上应用卷积神经网络进行分类。测试影像 1 及其对应结果如图 7 和表 2 所示。其中图 7(b) 真实类别图中的灰色部分为道路阴影, 其余地物类型注释与图 4(c) 中相同。表 2 中数字为分类结果中像素的统计数目。

在图 7 结果中看到, 黑色框中所示卷积神经网络在道路和灰色屋顶之中存在明显错分。在表 2 中也发现, 大量道路被错分为灰色屋顶。此外, 树木和草地之间也存在明显错分, 其余分类基本准确。

#### 实验2 卷积神经网络与SVM结合遥感影像分类



(a)滑窗取窗口分类结果 (b)SLIC取窗口分类结果 (c)Mean Shift取窗口分类结果

图6 训练数据中不同采样窗口影像的卷积神经网络分类结果

表1 滑动窗口法分类和本文选择基元分类实验精度对比

方法		分类类别									总体精度
		田地	树木	草地	荒地	道路	灰色屋顶	水体	珊瑚色屋顶	蓝色屋顶	
滑动取窗分类	生产者精度	93.54	90.13	55.98	—	70.17	93.96	91.82	93.59	—	88.97
	用户精度	95.28	87.54	60.08	0	91.52	75.59	93.97	85.92	0	
本文选取基元分类	生产者精度	96.36	86.47	49.55	—	67.83	92.42	95.47	91.64	—	88.99
	用户精度	93.98	87.18	66.10	0	97.36	74.96	88.49	87.77	0	

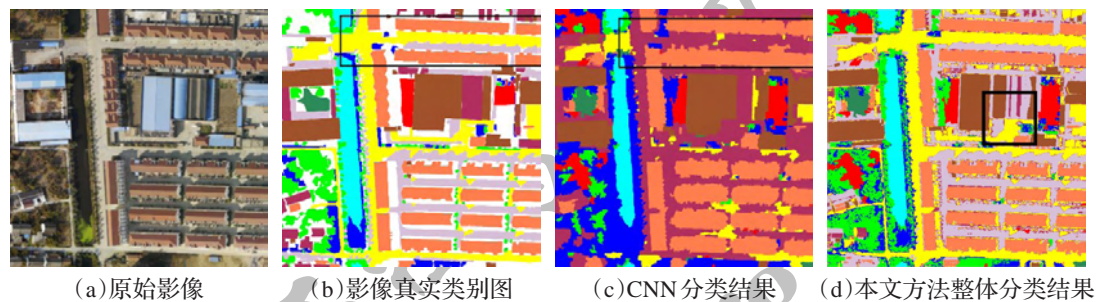


图7 测试影像1分类结果图

表2 测试影像1使用本文基元分类的分类结果混淆矩阵											像素
	田地	树木	草地	荒地	道路	灰顶	水体	珊瑚顶	蓝顶	阴影	用户精度/%
田地	18 080	9 042	1 168	217	441	0	0	0	502	3 309	55.19
树木	1	22 299	3 834	27	383	26	737	0	0	1 132	78.41
草地	2 486	64 067	46 902	0	15 407	4 121	9 252	558	649	21 035	28.52
荒地	0	1 204	768	8 141	617	0	0	2	0	2 630	60.93
道路	0	1 486	1 262	0	50 120	1 934	0	87	1 459	13 734	71.52
灰色屋顶	823	7 392	3 020	58	142 684	60 621	93	2 399	6 678	139 731	16.68
水体	0	955	829	0	0	0	61 397	0	106	0	97.01
珊瑚色屋顶	0	1 332	1 327	1 819	4 535	3 729	4	260 924	2 234	24 371	86.90
蓝色屋顶	902	1 040	385	0	1 340	3 968	1	2	138 745	6 023	91.04
阴影	0	0	0	0	0	0	0	0	0	0	—
生产者精度/%	81.11	20.49	78.83	79.33	23.25	81.48	85.89	98.85	92.27	0	—
总体精度/%	56.14										

表3 测试影像1在本文方法下的分类结果精度											%
精度	分类类别										总体精度
	田地	树木	草地	荒地	道路	灰色屋顶	水体	珊瑚色屋顶	蓝色屋顶	阴影	
生产者精度	81.11	61.31	44.32	79.33	90.95	25.67	85.89	93.30	87.13	79.86	79.29
用户精度	55.46	57.83	43.45	71.07	84.15	52.49	97.01	96.30	99.26	68.23	

针对卷积神经网络初始分类中的易错分类别,引入支撑向量机进行分类。图4(a)一些道路在房屋周围,从两种方法的混淆矩阵和目视分类结果中也可发现:(1)房屋周围的道路有很明显的错分。(2)一些树木被分类为草地,存在异常。(3)阴影主要集中在道路中。分析混淆矩阵的结果,并结合影像数据发现:房屋和道路在构造采样窗口影像时候在地物信息上有一定的相似性,因而不好区分。树木和草地在道路两旁同时出现,导致卷积神经网络无法对两类有效区分。因此,在影像数据使用SVM进一步提高精度,混合类别为道路-房屋-阴影、道路-树木-草地。实验中混合类别中每类随机选取1 000像素点,作为SVM的训练样本。分类结果如图7(d)所示,精度表如表3所示。

对比图7(c)和图7(d)结果发现,基于卷积神经网络分类结果基础上构建的混合类别中,仅使用光谱信息作为分类特征,即可分类道路和屋顶。但阴影和灰色屋顶中存在大量错分,说明使用光谱特征无法对灰色屋顶和阴影部分有效区分。影像中黑框部分所示,发现阴影区域与灰色屋顶部分特性相似,不易区分,未来考虑使用其他方法进行再分类。

实验3 本文方法与其他方法实验结果对比

为比较本文方法与常用分类方法的性能,使用常用方法与本文分类方法进行对比:基于光谱信息的SVM和基于光谱和dense SIFT的SVM。分类结果如图8所示,精度对比如图9所示。

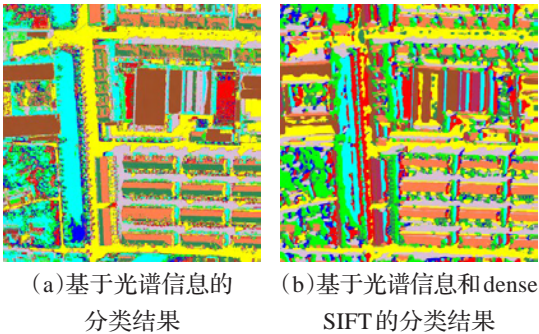


图8 常用方法分类结果

从图9的分类结果看,第二种对比方法水体的生产者精度比本文方法的精度高,但是用户精度比较低。生产者精度和用户精度分别对应漏分误差和错分误差,即在水体的分类结果中,对比算法的高生产者精度是以大量错分为基础的,这显然会影响其他类别的分类效果。



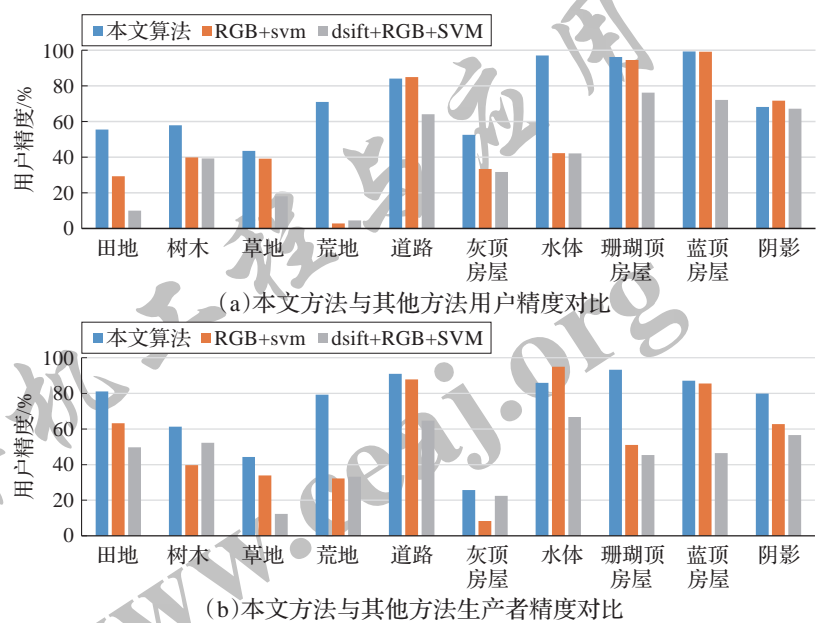


图9 本文方法与其他方法用户精度和生产者精度对比

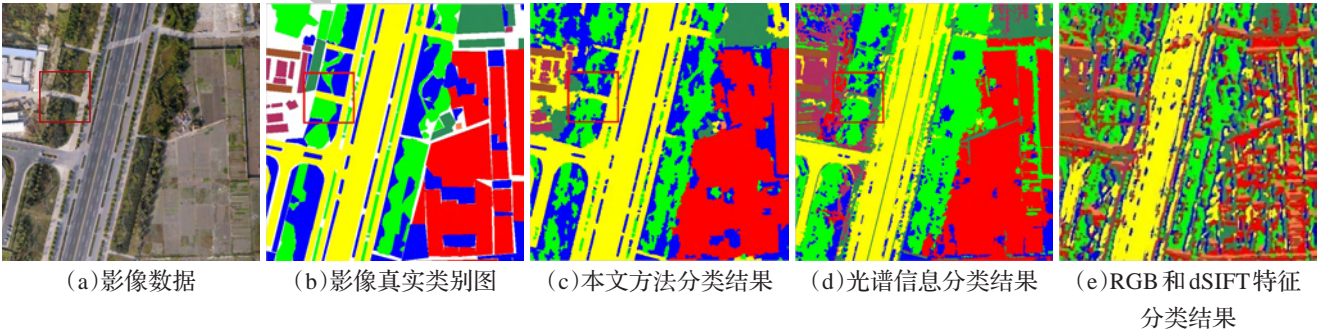


图10 测试影像二中本文方法和其他两种方法对比

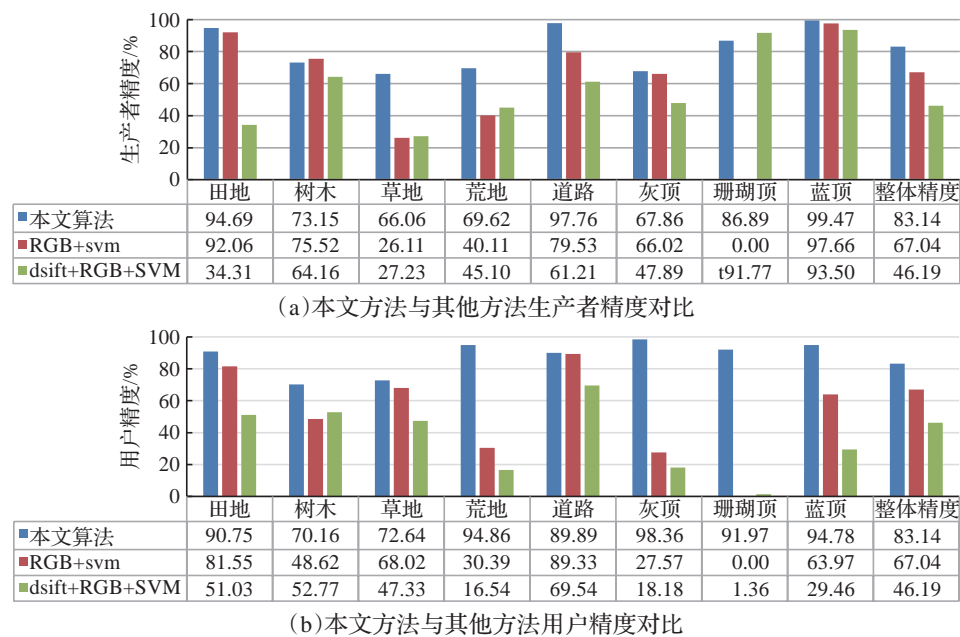


图11 测试影像二中本文方法与其他方法生产者精度和用户精度对比结果

本文的方法克服了上述问题。

本文另一遥感影像数据分类结果如图10所示。本文方法和其他方法对比如图11所示,图11中表格数据对应生产者精度和用户精度。从生产者精度和用户精度上看,本文方法优于其他两种方法,整体精度约有10%的提升。

#### 4 结束语

本文将卷积神经网络应用于遥感影像分类,在过分割的基础上确定了采样窗口影像,结合卷积神经网络实现了高分辨率遥感影像分类,并使用层次分类进一步提高了最终分类精度。实验结果表明,本文方法可提高卷积神经网络在高分辨率遥感影像中分类的效率。相比于传统人工设计的特征表示方法,基于卷积神经网络学习得到的特征具有更好的可区分性。下一步研究工作是结合卷积神经网络和影像中的结构信息得到更好的数据特性描述,提高高分辨率遥感影像分类精度。

#### 参考文献:

- [1] Xu Linlin, Li Jonathan. Bayesian classification of hyperspectral imagery based on probabilistic sparse representation and markov random field[J]. *Geoscience and Remote Sensing Letters*, 2014, 11(4): 823-827.
- [2] Shackelford A K, Davis C H. A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(9): 1920-1932.
- [3] Zhang L, Huang X, Huang B, et al. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(10).
- [4] Yuan J, Wang D L, Li R. Remote sensing image segmentation by combining spectral and texture features[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(1): 16-24.
- [5] Lowe D G. Object recognition from local scale-invariant features[C]// *Proceedings of the 7th IEEE International Conference on Computer Vision*, 1999: 1150-1157.
- [6] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]// *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005: 886-893.
- [7] Xu S, Fang T, Li D, et al. Object classification of aerial images with bag-of-visual words[J]. *Geoscience and Remote Sensing Letters*, 2010, 7(2): 366-370.
- [8] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521: 436-444.
- [9] Dos Santos J A, Gosselin P H, Philipp-Foliguet S, et al. Interactive multiscale classification of high-resolution remote sensing images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, 6(4): 2020-2034.
- [10] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619.
- [11] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]// *Advances in Neural Information Processing Systems*, 2012: 1097-1105.
- [12] Bouman C, Liu B. Multiple resolution segmentation of textured images[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(2): 99-113.
- [13] Pan S J, Yang Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [14] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 248-255.
- [15] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]// *Proceedings of the 22nd ACM International Conference on Multimedia*, 2014: 675-678.
- [16] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2274-2282.