



# 利用PSP Net实现湖北省遥感影像 土地覆盖分类

柴华彬<sup>1</sup> 严超<sup>1</sup> 邹友峰<sup>1</sup> 陈正超<sup>2</sup>

<sup>1</sup> 河南理工大学测绘与国土信息工程学院,河南 焦作,454003

<sup>2</sup> 中国科学院遥感与数字地球研究所,北京,100094

**摘要:**提出了一种基于金字塔场景解析网络(pyramid scene parsing net, PSP Net)的深度学习算法。以湖北省遥感影像为实验数据,借助PSP Net的上下文场景解析能力,研究湖北省30 m分辨率的土地覆盖。实验使用了湖北省Landsat卫星影像中507景900×600像素的标准分幅影像,通过预处理生成了适用于深度学习的样本集。选择其中300景为样本,包括训练集240个、预测集44个和验证集16个。使用快速特征嵌入卷积结构(convolutional architecture for fast feature embedding, CAFFE)下的PSP Net模型对样本数据进行训练,设置了 $10 \times 10^{-10}$ 的学习率,选择了第100万次的训练模型,很好地防止了数据的过拟合。通过模型的泛化和样本的泛化与迭代,对湖北省2000年、2005年、2010年3期的Landsat卫星影像土地覆盖进行分类,分类精度分别达到82.2%、83.4%和83.7%。研究结果表明,基于PSP Net的深度学习算法可以快速、有效和精确地实现大范围的遥感影像土地覆盖分类。

**关键词:**遥感影像;土地覆盖分类;深度学习;金字塔场景解析网络(PSP Net)

**中图分类号:**P237

**文献标志码:**A

遥感影像分类方法是研究土地覆盖及其变化的最重要手段。传统的遥感影像分类方法有目视解译法、统计分析法、聚类分析法等<sup>[1-3]</sup>。这些方法虽然简单、灵活,但是其局限性大且花费时间长、解译精度差,已经不适合当今的海量遥感影像分类。之后发展的支持向量机、决策树和人工神经网络等分类法<sup>[4-6]</sup>,在数据处理的速度上无法满足数据的大尺度和时效性。

近年来,深度学习的出现为遥感影像分类提供了新方法<sup>[7-9]</sup>。深度学习是一种深层次结构的神经网络,可以比支持向量机、决策树和人工神经网络等浅层结构模型更好地提取遥感影像的特征,在影像分类中,取得了比以往更高的精度,有效地推动了遥感影像自动化、智能化解译的发展。

在深度学习迅速发展之后,全卷积神经网络(fully convolutional network, FCN)使得图像分割进入了新的阶段,像素不仅能够按照物体类别进行分类,图像分割精度也得到提升<sup>[10-12]</sup>。目前,基于FCN的模型的主要问题是缺乏合适的策略来

利用全局场景类别线索,不能有效地处理场景之间的关系和全局信息。

金字塔场景解析网络(pyramid scene parsing net, PSP Net)是基于FCN的语义分割网络<sup>[13-16]</sup>。它可以将FCN的像素预测框架中嵌入困难场景的上下文特征,完整理解场景的进行,实现对每个像素类别、位置和形状的准确预测,并将局部和全局信息融合在一起,提出一个适度监督损失的优化策略。

随着国家经济的发展,各类建设用地迅速增多,城市面积急剧扩大,土地覆盖变化剧烈。Landsat 5是目前运行时间最长的对地观测卫星,它搭载的专题测图仪(thematic mapper, TM)为地球提供了近29年的成像数据。使用深度学习算法能够对TM数据进行语义分割,更好地实现遥感影像的分类。

本文使用507景900×600像素的湖北省标准分幅影像,借助PSP Net的上下文场景解析能力,通过模型的泛化和样本的泛化与迭代,获得了湖北省30 m分辨率的土地覆盖分类。

## 1 PSP Net 的构建

### 1.1 全卷积神经网络的描述

FCN 能够通过构建多层的卷积层自动提取图像上的特征。FCN 结构和其他神经网络结构类似<sup>[13,17]</sup>,分为输入层、隐藏层和输出层,如图 1 所示。输入层有  $n$  个节点,隐藏层有  $m$  个节点,输出层有  $q$  个节点,输入层和隐藏层间的权重为  $R_{ki}$ ,隐藏层和输出层间的权重为  $W_{jk}$ 。

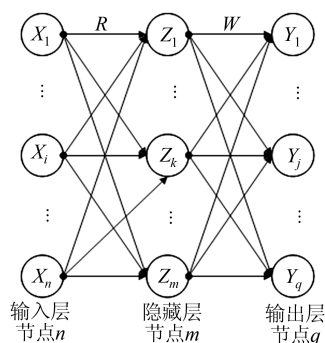


图 1 全卷积神经网络结构

Fig.1 Structure of Fully Convolutional Network

FCN 的隐藏层包含卷积层、激励函数和池化层。

1) 卷积层。该层的功能是对输入数据进行特征提取,其内部包含多个卷积核,组成卷积核的每个元素都对应一个权重系数和一个偏差量。卷积层内每个神经元都与前一层中位置接近的区域的多个神经元相连。以二维卷积核为例<sup>[18]</sup>,其模型为:

$$Z^{l+1}(i,j)=[Z^l \otimes w^{l+1}](i,j)+b \quad (1)$$

式中,  $(i,j) \in \{0,1,2 \dots L_{l+1}\}$ ,  $L_{l+1} = \frac{L_l + 2p - f}{S_0} + 1$ ;

$b$  为偏差量;  $Z^l$  和  $Z^{l+1}$  表示第  $l+1$  层的卷积输入和输出;  $L_{l+1}$  为  $Z_{l+1}$  的尺寸;  $Z(i,j)$  为对应特征图的像素;  $f$ 、 $S_0$  和  $p$  是卷积层参数,对应卷积核大小、步长和填充层数,共同决定卷积层输出特征图的尺寸。

2) 激励函数。类似于其他深度学习算法,激励函数操作通常在卷积核后,卷积层中包含激励函数以协助表达复杂特征。

设卷积层为第  $l$  层,令  $N^l$  表示  $l$  层的特征映射数量,每个特征映射被表示为  $h_j^l(j=1,2 \dots N^l)$ ,该卷积层被参数化为两个数组层中的第  $j$  个特征映射  $h_j^l$ ,每个卷积核充当特征检测器,用于通过卷积操作提取图像特定的特征。

为了获得  $h_j^l$ ,将输入的特征图  $h_j^{l-1}(j=1,2 \dots$

$N^{l-1})$  与相应的滤波器  $w_{ij}^l$  进行卷积,通过修正线性单元(rectified linear unit, ReLU)实现梯度下降及反向传播<sup>[15]</sup>,将第  $l$  层的特征图表示如下:

$$h_j^l = \Phi \left( \sum_{i=1}^{N^{l-1}} h_i^{l-1} \times w_{ij}^l + b_j^l \right) \quad (2)$$

3) 池化层。在卷积层进行特征提取后,输出的特征图被传递至池化层进行特征选择和信息过滤。池化层为采样特征映射层,通过降低特征的分辨率来防止过拟合。同时,能实现减少网络参数,保持图像的平移、伸缩和旋转不产生变形。本文选用的池化模型<sup>[19]</sup>为:

$$A_k^l(i,j) = \left[ \sum_{x=1}^f \sum_{y=1}^f A_k^l(S_0 i + x, S_0 j + y)^p \right]^{1/p} \quad (3)$$

式中,步长  $S_0$ 、像素  $(i,j)$  的含义与卷积层相同;  $p$  为预指定参数。当  $p=1$  时,在池化区域内取均值;当  $p \rightarrow \infty$  时,在池化区域内取极大值。

### 1.2 金字塔池化结构的引入

FCN 最主要的优势是利用跳跃结构对不同池化层的结果进行采样,再通过优化这些采样结果,得到精确分割结果。对具体目标检测或者分割任务来说,FCN 网络本身具有强大的特征提取能力,但是其分割结果比较粗糙。

利用金字塔池化结构代替 FCN 中最后一层池化层,能够解决输入网络中图片尺寸固定的限制,同时通过金字塔池化层,把上一层卷积层的特征图以不同尺度进行特征池化,能增加算法的鲁棒性,提高识别精度。

### 1.3 CAFFE 深度学习框架的搭建

快速特征嵌入卷积结构(convolutional architecture for fast feature embedding, CAFFE)是一个兼具表达性、速度和思维模块化的深度学习框架,支持多种类型的深度学习架构,用于图像的分类和分割<sup>[20-21]</sup>。它是开源的框架,核心语言是 C++,支持命令行、Python 和 Matlab 接口,既可以在中央处理器上运行,也可以在图形处理器上运行。

首先,基础层经过 CAFFE 深度学习框架的预训练模型和全卷积策略提取特征图层,提取后的特征图层是输入影像的 1/8 大小,特征图层经过金字塔池化模块得到融合的带有上下文整体信息的特征;然后,将池化前的特征图层与采样结果相连接;最后,经过卷积层后得到最终输出结果。



## 2 数据来源与实验方法

### 2.1 数据来源与样本选择

研究区影像是湖北省2000年、2005年、2010年3期Landsat 5 TM影像的6波段数据,这些数据来源于地理空间数据云和美国国家地质调查局网站。所选影像的成像时间多为夏季,且云量少。对影像进行辐射定标、大气校正、镶嵌及裁剪等预处理操作后分幅导出,得到507景 $900 \times 600$ 像素的影像,样本数据同样地将分类结果通过程序规整后,再分幅导出得到与遥感影像名称相对应的507景分类结果。

样本选择湖北省的土地覆盖数据,采用面向对象解译方法完成。根据土地覆盖类型的特征与光谱规律<sup>[12]</sup>,参数层划分为4个层次与5个分支节点,具体节点为水面与非水面、植被与非植被、线性与非线性、耕地与非耕地、落叶与非落叶。下层依据区域特征进一步设计,通过对象的解译标志库和样本训练,建立分类决策树指标与决策树结构,通过决策树的分级,不断提纯类型,得到单个类别划分的结果<sup>[22-23]</sup>。

通过交互式数据语言(interactive data language, IDL)程序,重构原始样本中38个二级分类结果,得到包括未分类的8个类别,获得的样本分类体系如表1所示<sup>[24]</sup>。

表1 样本分类体系

Tab.1 Sample Classification System

类别	二级分类结果
林地	常绿阔叶林、落叶阔叶林、常绿针叶林、落叶针叶林、针阔混交林、常绿阔叶灌木林、落叶阔叶灌木林、常绿针叶灌木林、乔木园地、灌木园地、乔木绿地、灌木绿地等
草地	草甸、草原、草丛、草本绿地等
湿地	森林湿地、灌丛湿地、草本湿地等
水体	湖泊、水库/坑塘、河流、运河、水渠等
耕地	水田、旱地等
人工地表	居住地、工业用地、交通用地等
裸地	苔藓、地衣、裸岩、裸土、沙漠、沙地、盐碱地等

### 2.2 软硬件准备及框架搭建

本文采用Ubuntu16.04.4 LTS操作系统,在ArcGIS、ENVI、Erdas等专业软件以及IDL语言和Python语言编程技术的支持下,分析了湖北省土地覆盖类型。搭建实验使用的深度学习CAFFE框架需要安装依赖包,依次安装英伟达显卡驱动、通用并行计算架构(compute unified device architecture, CUDA)、深度神经网络(com-

pute unified deep neural network, CuDNN)库和开放源代码计算机视觉(open source computer vision, OpenCV)类库,将PSP Net源码编译CAFFE框架,完成深度学习框架的搭建。

### 2.3 实验方法及流程

利用ENVI 5.3等软件对源数据进行辐射定标和大气校正,再利用像素工厂进行影像拼接与裁剪。将通过预处理得到的数据与样本数据生成影像样本集。

使用CAFFE深度学习框架下的PSP Net模型对湖北省Landsat TM影像的样本数据进行训练,对模型精度评价,通过模型的泛化和样本的泛化与迭代,实现湖北省遥感影像土地覆盖分类。

实验选出的样本分为训练集、预测集和验证集。通过编写程序,将样本的最后一个波段叠加到影像的最后一个波段,使TM影像变成7波段数据,其中240个 $900 \times 600$ 像素样本作为训练集,44个样本作为测试集,16个样本作为验证集,如图2所示。

## 3 数据拟合与模型精度评价

对利用深度学习PSP Net语义分割生成的3期湖北省土地覆盖图进行精度评价。在统计学中,假设数据满足独立同分布,可以用已有数据进行训练,使用训练所得到的模型拟合未来数据。但是现实中独立同分布这种假设是不成立的,数据分布可能会产生变化,并且当前数据量过少,不足以对整体数据集进行分布估计。为了提高预测模型的泛化能力,需要防止模型过拟合。

### 3.1 模型训练及选择

将2010年湖北省遥感影像数据集导入深度学习服务器,使用CAFFE深度学习框架中的Python程序进行运算。运算网络通过最核心的结构(金字塔池化模块)获得全局的特征信息,完成了数据集的生成。通过测试,设置 $10 \times 10^{-10}$ 的学习率,选择每100万次迭代运算获得的模型,对整个湖北省507景6波段TM影像数据进行预测,最终获得2010年期湖北省土地覆盖的预测结果,如图3所示。

为了评价预测模型的精度,采用总体精度指标(overall accuracy, OA)作为初步评价模型的预测精度指标<sup>[18]</sup>。OA表示被正确分类的样本数与研究区域总样本数之比,即

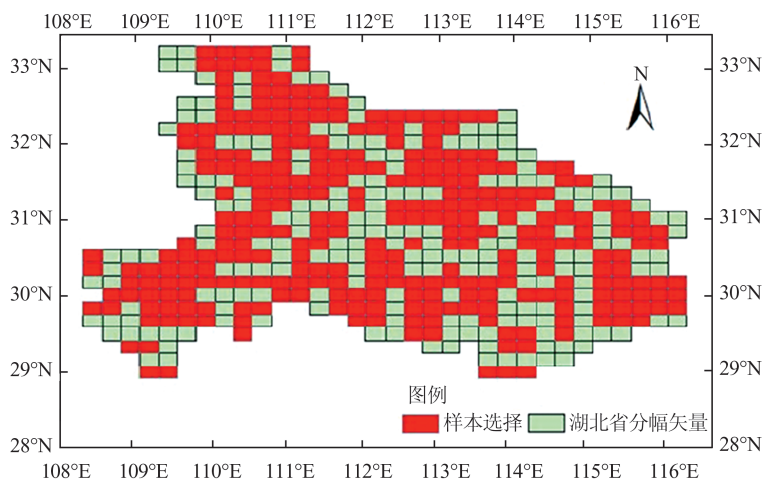


图 2 实验样本选择

Fig.2 Experimental Sample Selection

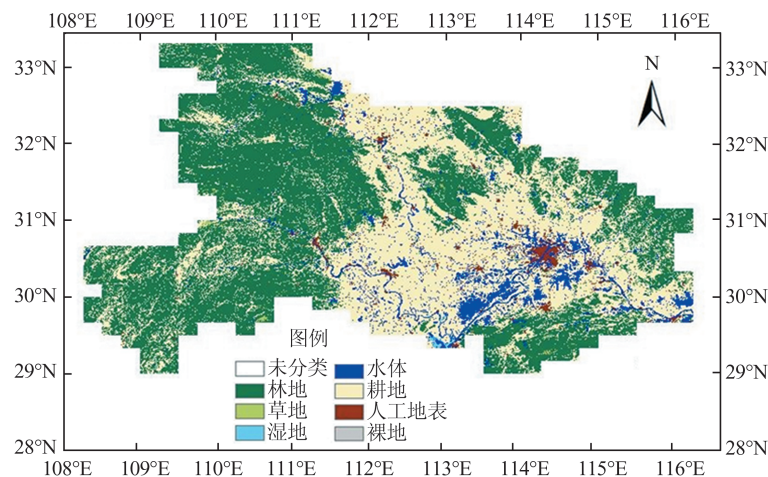


图 3 2010年湖北省土地覆盖分类结果

Fig.3 Result of Land Cover Classification About Hubei Province in 2010

$$OA = \frac{\sum_{i=1}^8 a_{ii}}{\sum_{i=1}^8 \sum_{j=1}^8 a_{ij}} \quad (4)$$

式中,  $i$  表示样本类别;  $a_{ii}$  表示第  $i$  类正确分类到  $i$  类的像元个数;  $a_{ij}$  表示把第  $i$  类分类到第  $j$  类的像元个数。

OA 的数值越大, 表示分类正确的结果越多, 分类精度越高, 结果越可信; 反之, 表示分类精度越低, 结果越不可信<sup>[25-26]</sup>。

本次实验在 507 景样本和预测数据中使用固定和随机两种方式, 在每景影像中选择 20 000 个同位点, 整个湖北省选择了 10 140 000 个点, 将这些点作为验证的总样本。通过波段运算, 获得一景样本和预测结果的差值文件。

### 3.2 模型的泛化及精度评定

模型的泛化是通过机器学习算法对具有同

一规律的学习集以外的数据, 经过训练网络能给出合适的输出结果。模型的泛化能力是指对于已知的数据(训练集)性能表现良好, 对于未知的数据(测试集)也应该表现良好。测试集的误差被称为泛化误差。模型的过拟合与欠拟合可以最直观地区分出模型泛化能力的优劣。

通过训练和选择获得 2010 年 OA 精度为 83.7% 的预测模型, 利用该模型的泛化对 2005 年数据的预测结果和 2005 年 TM 影像在 ArcGIS 中叠加后显示, 将透明度设置为 50%, 分类结果显示如图 4 所示。

### 3.3 样本的泛化及精度评定

虽然 PSP Net 网络对于错误的样本有一定的抗干扰的能力, 但是 2005—2010 年很多地方的变化很大, 若完全使用同一套样本数据, 则这些错误的样本会使模型快速过拟合, 精度会偏低。

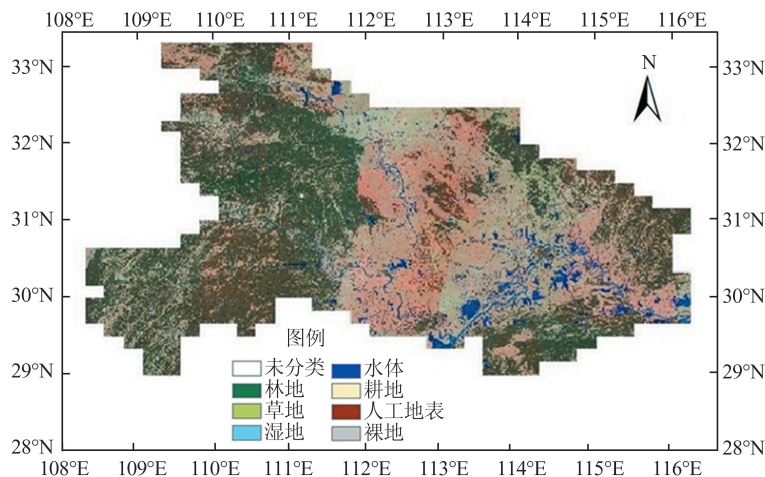


图4 2005年预测数据与影像叠加

Fig.4 Prediction Data of 2005 and Image Overlay

将2010年变化大的区域(如人工地表)使用矢量栅格化的工具将其标记并替换出来,与2010年的原始样本、分类结果较好的样本(如水体)和部分变化不大的样本(如大面积的林地、耕地、水体、草地和湿地等)合起来,组成整个原始的样本库。为了充分提高模型的泛化能力,需要对样本进行泛化和迭代,即不完全使用2010年的数据样本作为2005年数据样本进行深度学习和训练。

本次实验中,选择了100景来自模型泛化后得到的数据和100景2010年中变化不大的样本

数据组成了原始的样本库。将这些数据中的180景数据作为训练集,20景作为验证集和预测集。对于2005年的数据,初次训练时样本库的样本数量为200景。将这些样本和2005年的影像导入深度学习服务器,进行模型训练。

通过实验可知,训练次数设为100万次,可有效地防止数据过拟合。第1次迭代后,增加了20景样本进入训练集;第2次迭代后,增加了31景样本进入样本库;经过3次迭代后,对湖北省2005年TM影像进行预测,获得了较好的预测结果,如图5所示。

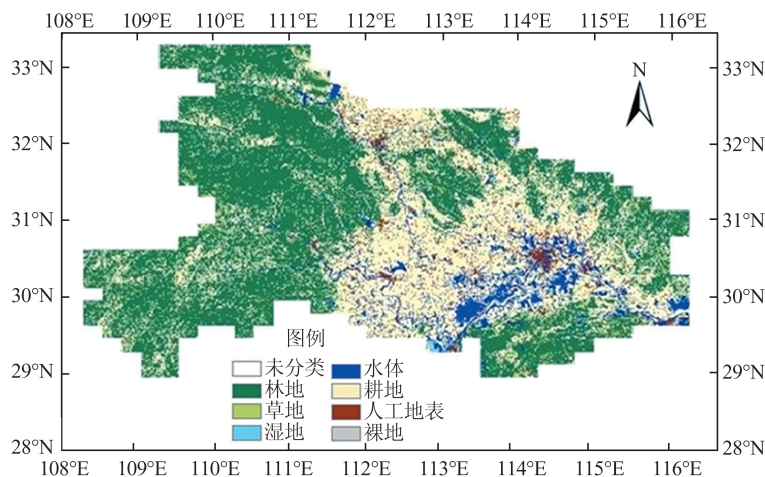


图5 2005年湖北省土地覆盖分类结果

Fig.5 Result of Land Cover Classification About Hubei Province in 2005

为了评价湖北省2005年土地覆盖的分类精度,随机且类别均匀地在该期影像及预测结果中找了1000个点,其中,有834个点的分类结果是正确的,根据OA评价出整个2005年预测结果的精度为83.4%。

使用同样的方法,获得了2000年TM影像的分类结果,如图6所示。

在2000年分类结果和影像中分别随机均匀选择了1000个点,其中,有822个点的分类结果正确,根据OA评价出整个2000年预测结果的精



度为 82.2%。

### 3.4 实验结果对比分析

利用基于 PSP Net 的深度学习算法与传统的

机器学习算法对遥感影像解译效果进行对比分析,从解译结果中分别选取两部分进行细节对比,如图 7 所示。

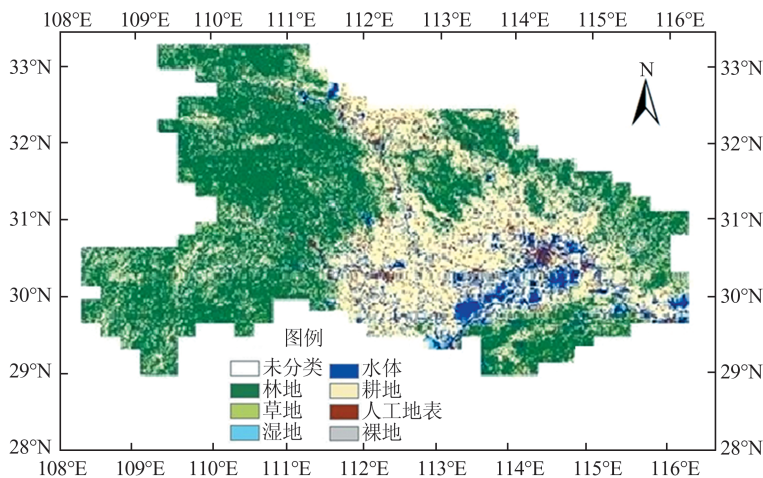


图 6 2000 年湖北省土地覆盖分类结果

Fig.6 Result of Land Cover Classification About Hubei Province in 2000

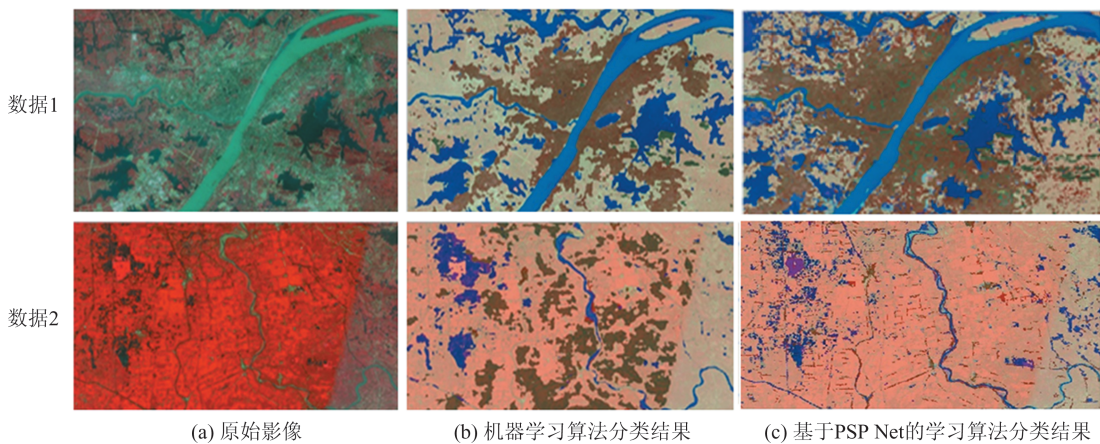


图 7 不同方法分类结果对比

Fig.7 Comparisons of Classification Results with Different Methods

数据 1 为人工地表聚集的武汉市区域。在传统的机器学习算法分类结果中,一些地类没有很好分出来,如城市区域应属于人工地表的类别,但是这些地方没有被模型区分出来,如图 7(b)所示。而利用基于 PSP Net 的深度学习算法分类结果中,人工地表、耕地、水体和林地等各个地类在武汉市区域都能较好地地区分出来,如图 7(c)所示。

数据 2 的中间部分明显是耕地,但是传统的机器学习算法将其识别为了林地,如图 7(b)所示。而基于 PSP Net 的深度学习算法能够将一些线状、面积很小的水体准确地区分出来,如图 7(c)所示。

## 4 结 语

本文针对湖北省 2000 年、2005 年和 2010 年 3 期 TM 影像的成像特点和波段特征,在 FCN 的基础上,提出了基于 PSP Net 的深度学习算法。通过修改和优化网络,排除过拟合效应后,使用 CAFFE 框架下的 PSP Net 模型对样本数据进行训练,在学习率为  $10 \times 10^{-10}$  的条件下,选择迭代次数为 100 万次的结果,有效地防止了数据的过拟合。

通过模型的泛化和样本的泛化与迭代,对湖北省 3 期的 TM 影像进行语义分割和土地覆盖分类,分类的总体精度分别达到 82.2%、83.4% 和 83.7%。研究表明,基于 PSP Net 的深度学习算

法能够快速、有效和精准地对大范围的土地覆盖进行分类。

### 参 考 文 献

- [1] Li Zhen, Zhang Yan, Yang Song, et al. Error Assessment of Extracting Morphological Parameters of Bank Gullies by Manual Visual Interpretation Based on QuickBird Imagery [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30(20): 179-186 (李镇, 张岩, 杨松, 等. QuickBird影像目视解译法提取切沟形态参数的精度分析[J]. 农业工程学报, 2014, 30(20): 179-186)
- [2] Wang Zhijie. Rapid Detection Method for Land Use Change Based on Remote Sensing Images Segmentation Units [J]. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, 2015, 39(3): 1-5 (王志杰. 基于遥感影像分割单元的土地利用变化快速检测方法[J]. 南京林业大学学报(自然科学版), 2015, 39(3): 1-5)
- [3] Huang Liwen, Mao Zhengyuan, Li Erzhen, et al. The Cluster Analysis Approaches Based on Geometric Probability and Its Application in the Classification of Remotely Sensed Images [J]. *Journal of Image and Graphics*, 2007, 12(4): 633-639 (黄利文, 毛政元, 李二振, 等. 基于几何概率的聚类分析方法及其在遥感影像分类中的应用[J]. 中国图像图形学报, 2007, 12(4): 633-639)
- [4] Tang Yunwei, Zhang Jingxiong. Land Cover Classification of Remotely Sensed Imagery Using Multiple-Point Geostatistics [J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(5): 546-550 (唐韵玮, 张景雄. 遥感影像土地覆盖分类的多点地统计学方法[J]. 武汉大学学报·信息科学版, 2014, 39(5): 546-550)
- [5] Lin Yi, Liu Bing, Chen Yingying, et al. Change Detection Method Based on Multi-feature Different Kernel SVM for Remote Sensing Imagery [J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(8): 978-982 (林怡, 刘冰, 陈映鹰, 等. 多特征差分核支持向量机遥感影像变化检测方法[J]. 武汉大学学报·信息科学版, 2013, 38(8): 978-982)
- [6] Zhang Xiaohe, Zhai Liang, Zhang Jixian, et al. Application of AdaTree Algorithm to Remote Sensing Image Classification [J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(12): 1460-1463 (张晓贺, 翟亮, 张继贤, 等. AdaTree算法在遥感影像分类中的应用[J]. 武汉大学学报·信息科学版, 2013, 38(12): 1460-1463)
- [7] Turner D P, Ritts W D, Cohen W B, et al. Scaling Gross Primary Production (GPP) Over Boreal and Deciduous Forest Landscapes in Support of MODIS GPP Product Validation [J]. *Remote Sensing of Environment*, 2003, 88(3): 256-270
- [8] Meng Qingxiang, Duan Xuelin. High Resolution Remote Sensing Image Scene Classification Based on DCNN [J]. *Journal of Central China Normal University (Natural Sciences)*, 2019, 53(4): 568-574 (孟庆祥, 段学琳. 基于DCNN的高分辨率遥感影像场景分类[J]. 华中师范大学学报(自然科学版), 2019, 53(4): 568-574)
- [9] Zhang Jia, Xie Yufeng. Summary of Hyperspectral Remote Sensing Image Classification Methods [J]. *Anhui Agricultural Science Bulletin*, 2017, 23(14): 155-165 (张佳, 谢玉凤. 高光谱遥感影像分类方法综述[J]. 安徽农学通报, 2017, 23(14): 155-165)
- [10] Wang Bin, Fan Donglin. Research Progress of Deep Learning in Classification and Recognition of Remote Sensing Images [J]. *Bulletin of Surveying and Mapping*, 2019(2): 99-102 (王斌, 范冬林. 深度学习在遥感影像分类与识别中的研究进展综述[J]. 测绘通报, 2019(2): 99-102)
- [11] Bhosle K, Musande V. Evaluation of Deep Learning CNN Model for Land Use Land Cover Classification and Crop Identification Using Hyperspectral Remote Sensing Images [J]. *Journal of the Indian Society of Remote Sensing*, 2019, 47(11): 1949-1958
- [12] Boualleg Y, Farah M, Farah I R. Remote Sensing Scene Classification Using Convolutional Features and Deep Forest Classifier [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(12): 1944-1948
- [13] Shi Binbin, He Haiqing, You Qi. A Method of Multi-scale Total Convolution Network Driven Remote Sensing Image Repair [J]. *Journal of Geomatics*, 2018, 43(3): 124-126 (石斌斌, 何海清, 游琦. 一种多尺度全卷积神经网络驱动的遥感影像修复方法[J]. 测绘地理信息, 2018, 43(3): 124-126)
- [14] Fu Weifeng, Zou Weibao. Review of Remote Sensing Image Classification Based on Deep Learning [J]. *Application Research of Computers*, 2018, 35(12): 521-524 (付伟峰, 邹维宝. 深度学习在遥感影像分类中的研究进展[J]. 计算机应用研究, 2018, 35(12): 521-524)
- [15] Zhang Chunlei. Application of Target Detection Technology Based on Military Aircraft Image Combined with FCN [J]. *Theory and Algorithm*, 2019(10): 53-55 (张春雷. 基于军事飞机图像结合FCN的目标检测技术应用[J]. 理论与算法, 2019(10): 53-55)
- [16] Yang Jianyu, Zhou Zhenxu, Du Zhenrong, et al.

- Rural Construction Land Extraction from High Spatial Resolution Remote Sensing Image Based on SegNet Semantic Segmentation Model[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(5): 251-258 (杨建宇, 周振旭, 杜贞容, 等. 基于 SegNet 语义模型的高分辨率遥感影像农村建设用地提取[J]. 农业工程学报, 2019, 35(5): 251-258)
- [17] Zhang Bing. Remote Sensing Big Data Era and Intelligent Information Extraction[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1 861-1 871 (张兵. 遥感大数据时代与智能信息提取[J]. 武汉大学学报·信息科学版, 2018, 43(12): 1 861-1 871)
- [18] Cai Bowen, Wang Shugen, Wang Lei, et al. Extraction of Urban Impervious Surface from High-Resolution Remote Sensing Imagery Based on Deep Learning [J]. *Journal of Geo-information Science*, 2019, 21(9): 1 420-1 429 (蔡博文, 王树根, 王磊, 等. 基于深度学习模型的城市高分辨率遥感影像不透水面提取[J]. 地球信息科学学报, 2019, 21(9): 1 420-1 429)
- [19] Liu Wanjun, Liang Xuejian, Qu Haicheng. Learning Performance of Convolutional Neural Networks with Different Pooling Models[J]. *Journal of Image and Graphics*, 2016, 21(9): 1 178-1 190 (刘万军, 梁雪剑, 曲海成. 不同池化模型的卷积神经网络学习性能研究[J]. 中国图象图形学报, 2016, 21(9): 1 178-1 190)
- [20] Takeki A, Trinh T T, Yoshihashi R, et al. Combining Deep Features for Object Detection at Various Scales: Finding Small Birds in Landscape Images [J]. *IPSJ Transactions on Computer Vision and Applications*, 2016, 8(1): 5-7
- [21] Feng Jin, Li Yuhui. Synthetic Recognition of Vehicle Color Based on Convolution Neural Network [J]. *Electronic Science and Technology*, 2018, 31(6): 89-92 (冯锦, 李玉惠. 基于卷积神经网络的车型颜色综合识别[J]. 电子科技, 2018, 31(6): 89-92)
- [22] Zhao Zhongming, Gao Lianru, Chen Dong, et al. Satellite Remote Sensing and Image Processing Platform Development [J]. *Journal of Image and Graphics*, 2019, 24(12): 2 098-2 110 (赵忠明, 高连如, 陈东, 等. 卫星遥感及图像处理平台发展[J]. 中国图象图形学报, 2019, 24(12): 2 098-2 110)
- [23] Liu Haiying, Gao Jixi, Li Zhenghai. The Advances in the Application of Remote Sensing Technology to the Study of Land Covering and Land Utilization[J]. *Remote Sensing for Land & Resources*, 2001, 13(4): 7-12 (柳海鹰, 高吉喜, 李政海. 土地覆盖及土地利用遥感研究进展[J]. 国土资源遥感, 2001, 13(4): 7-12)
- [24] Ji Shunping, Tian Siqi, Zhang Chi. Urban Land Cover Classification and Change Detection Using Fully Atrous Convolutional Neural Network [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(2): 233-241 (季顺平, 田思琦, 张驰. 利用全空洞卷积神经网络进行城市土地覆盖分类与变化检测[J]. 武汉大学学报·信息科学版, 2020, 45(2): 233-241)
- [25] Zhang Ce, Zang Shuying, Jin Zhu, et al. Research on Zhalong Wetland Classification Based on Support Vector Machines [J]. *Wetland Science*, 2011, 9(3): 263-269 (张策, 臧淑英, 金竺, 等. 基于支持向量机的扎龙湿地遥感分类研究[J]. 湿地科学, 2011, 9(3): 263-269)
- [26] Wang Haijun. The Application Research of AGA-BP Model in Remote Sensing Image Classification [J]. *Computer Measurement and Control*, 2017, 25(5): 212-214 (王海军. AGA-BP 模型在遥感影像分类中的应用研究[J]. 计算机测量与控制, 2017, 25(5): 212-214)

## Land Cover Classification of Remote Sensing Image of Hubei Province by Using PSP Net

CHAI Huabin<sup>1</sup> YAN Chao<sup>1</sup> ZOU Youfeng<sup>1</sup> CHEN Zhengchao<sup>2</sup>

<sup>1</sup> School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China

<sup>2</sup> Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China

**Abstract: Objectives:** Pyramid scene parsing net (PSP Net), a kind of neural network with deep structure, extracts the features of remote sensing image better than the traditional model, e.g. artificial neural network, and supports vector machine. This algorithm can embed the contextual features of difficult scenes in-



to the pixel prediction framework of the full convolutional neural network (FCN), fully understand the scene, realize accurate prediction of each pixel category, location and shape, and fuse local and global information together to propose an optimization strategy for moderate supervised loss. A deep learning algorithm based on PSP Net is proposed to achieve higher accuracy in image classification and effectively promote remote sensing image automation and intelligent interpretation. **Methods:** Conventional architecture for fast feature embedding (CAFFE) is a deep learning framework with is expressive, fast and thought modular, and supports a variety of deep learning architectures for image classification and segmentation. PSP Net model under the CAFFE framework is used by modifying and optimizing the network to eliminate the overfitting effect. Using the remote sensing image of Hubei Province as the experimental data, the land with the 30 m resolution is studied with the aid of the analysis ability of the context scene of PSP Net. The Python program is used in the CAFFE depth learning framework for operation. The operation network obtains the global feature information through the core structure (pyramid pooling module) and completes the generation of the data set. In the experiment, the 507 standard partial images of  $900 \times 600$  pixels in the landsat image of Hubei Province are used, and the sample sets suitable for depth learning are generated by pre-processing. Among these, 300 data are selected, including 240 training sets, 44 prediction sets, and 16 validation sets. The overall accuracy (OA) index is used as the prediction accuracy index of the preliminary evaluation model to evaluate the accuracy of the prediction model. **Results:** PSP Net model under the deep learning framework of CAFFE is used to train the sample data. The learning rate of  $10 \times 10^{-10}$  is set, and the million times training model is selected to protect against overfitting of the data. Through the generalization of the model and the generalization and iteration of the samples, the land cover of the 3 phases of TM images about Hubei Province are classified, and the classification accuracy is 82.2%, 83.4% and 83.7%, respectively. **Conclusions:** The results show that the land cover classification of remote sensing images can be realized quickly, effectively and accurately by PSP Net.

**Key words:** remote sensing image ; land cover classification ; deep learning ; pyramid scene analysis network (PSP Net)

**First author:** CHAI Huabin, PhD, professor, specializes in the deformation monitoring and remote sensing image processing. E-mail: chaih@hpu.edu.cn

**Foundation support:** The National Natural Science Foundation of China(U1810203).

**引文格式:** CHAI Huabin, YAN Chao, ZOU Youfeng, et al. Land Cover Classification of Remote Sensing Image of Hubei Province by Using PSP Net[J]. Geomatics and Information Science of Wuhan University, 2021, 46(8): 1224-1232. DOI:10.13203/j.whugis20190296(柴华彬, 严超, 邹友峰, 等. 利用PSP Net实现湖北省遥感影像土地覆盖分类[J]. 武汉大学学报·信息科学版, 2021, 46(8): 1224-1232. DOI:10.13203/j.whugis20190296)