

DOI:10.13332/j.1000-1522.20180342

## 基于卷积神经网络的高分二号影像林分类型分类

江 涛 王新杰

(北京林业大学林学院 北京 100083)

**摘要:**【目的】基于遥感影像的林分类型分类在现代林业中是一项重要的应用。本文试图构建一个基于高分二号(GF-2)影像林分类型分类的 CNN 模型,探索 CNN 在遥感图像像素级分类这一领域的发展潜力。【方法】以 GF-2 卫星遥感影像为数据源,利用 Tensorflow(一种开源用于机器学习的框架)构建 4 种不同图像斑块大小( $m = 5, 7, 9, 11$ )为输入的 CNN,同时以传统的神经网络模型,多层感知器(multi-layer perceptron, MLP)为基准,比较不同图像斑块大小下的 CNN 分类图的分类效果和分类精度。【结果】实验分类结果表明:CNN( $m = 9$ )得出最高的分类精确度,总体精度比 MLP 和 CNN( $m = 5, 7, 11$ )分别高出 10.91% 和 6.55%、1.3%、2.54%。分类图的可视化结果也表明 CNN( $m = 9$ )更好地解决了“椒盐现象”与过度平滑后的边界不确定性问题。【结论】CNN 能够在利用高分影像光谱特征的同时充分挖掘影像的空间特征,从而提高分类精度,同时在利用 CNN 基于遥感影像分类时,根据数据源以及地物的特点选择合适的图像斑块大小是提高分类精度与分类效果的关键措施。

**关键词:** CNN; MLP; GF2; 林分类型; 图像斑块大小

**中图分类号:** S771.8 **文献标志码:** A **文章编号:** 1000-1522-(2019)09-0001-10

**引文格式:** 江涛,王新杰. 基于卷积神经网络的高分二号影像林分类型分类[J]. 北京林业大学学报, 2019, 41(9):1-10.  
JIANG Tao, WANG Xinjie. Convolutional neural network for GF-2 images stand type classification[J]. Journal of Beijing Forestry University, 2019, 41(9): 1-10.

## Convolutional neural network for GF-2 images stand type classification

JIANG Tao WANG Xinjie

(College of forestry, Beijing Forestry University, Beijing 100083, China)

**Abstract:** [Objective] The classification of stand type based on remote sensing imagery is an important application in modern forest. In recent years, many studies have explored this territory using multiple data sources and classification algorithm. Convolutional neural network (CNN), a new neural network algorithm, has higher accuracy in pattern recognition, scene classification and objective detection because of its unique structure and deep learning technology. The purpose of this paper is to propose a convolutional neural network system tailored for GF-2 (a high-resolution multispectral remote sensing data) applied in stand type classification on pixel-level. [Method] We chose different image patches size (i.e.  $m = 5, 7, 9$  and 11) for building CNN and multilayer perceptrons (MLP) as benchmark in tensorflow (an open source library of machine intelligence), to train and compare classification accuracy of model. [Results] Experimental results showed that the CNN ( $m = 9$ ) outperformed MLP and CNNs ( $m = 5, 7$  and 11) by 10.91%, 6.55%, 1.3% and 2.54%, respectively, in overall classification accuracy. And the CNN ( $m = 9$ ) alleviates the effect of salt-and-pepper and boundary uncertainties greatly in visual assessment. [Conclusions] CNN can fully exploit the spatial features of images while utilizing the spectral features of high-resolution images to improve classification accuracy. And on remote sensing image classification based on CNN, selecting the appropriate

收稿日期: 2018-10-19 修回日期: 2019-07-05

基金项目: 国家重点研发计划项目(2017YFC050410101), 中央高校基本科研业务费专项(BLJD200907、JD2010-2)。

第一作者: 江涛, 主要研究方向: 林业遥感与信息技术。Email: taoorwell@bjfu.edu.cn 地址: 100083 北京市海淀区清华东路 35 号北京林业大学林学院。

责任作者: 王新杰, 教授, 主要研究方向: 森林资源监测和林业遥感。Email: xinjiew@bjfu.edu.cn 地址: 同上。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

image patch size according to the data source and the features of the objective is the key measure to improve the classification accuracy and classification effect.

**Key words:** CNN; MLP; GF-2; stand type; image patches size

林分类型是森林资源调查中一个重要的调查因子,其对森林的经营管理与生态保护起到重要作用。林分中树种的组成与分布将直接影响森林的生长与收获,对不同树种组成的林分采取的经营措施也不相同。在生态保护问题当中,了解森林的树种多样性,也有助于理解各物种在生态系统中起到的生态作用。通过实地踏查获取大区域的树种组成信息成本较高,周期较长,遥感数据能够及时有效的获取大区域的地面信息,并且随着遥感技术的发展,遥感影像在空间、光谱、时间分辨率上都有巨大的提升,借助遥感手段获取林分的类型以及观测森林面积的消长已经成为一个重要的研究方向。

高分二号卫星是我国高分系列卫星中首颗空间分辨率优于1 m的光学遥感卫星,也是我国目前分辨率最高的民用陆地观测卫星,由于高分二号在有较多地物光谱信息的同时包含丰富的空间纹理信息,高分二号卫星影像被广泛的应用在土地利用调查、矿产资源开发、农作物产量及生长监测和森林资源等任务当中。不少学者在对高分二号影像预处理方面,包括正射矫正和图像融合等技术做出了大量的研究<sup>[1-2]</sup>,这为推动高分二号影像的应用做出了很大的贡献。

在影响影像分类精度研究中,数据源的重要性不言而喻,但在另一方面,选择合适的分类算法也显得尤为关键。树种分类问题在本质上是属于分类问题,在早期的分类研究当中,一些分类算法例如监督分类中的最大释然法(maximum likelihood classification),非监督分类中的聚类算法(K-means)被广泛的应用于分类任务当中。但在近十几年,随着计算机技术的发展,一些新兴的分类算法,例如支持向量机(support vector machine)<sup>[3]</sup>、随机森林(random forest)<sup>[4]</sup>和神经网络,这些非参数的机器学习分类算法相较于传统的分类方法取得较好的分类效果,其中由于神经网络模型的参数多,数据拟合程度高,模型可调控性好等特点,对神经网络模型的研究一直在持续。神经网络模型是受到人类大脑的运作机制而启发的,在人的视觉大脑皮层中,包含数以百万的神经元,神经元之间的轴突数以亿计,因此能够处理复杂的图像识别任务<sup>[5]</sup>。多层感知器(multi-layer perceptron)是一种典型的神经网络模型,被设计用来处理复杂的分类与回归预测任务。在早期的一些研究当中,MLP在基于遥

感影像的土地利用与土地分类任务中取得良好的效果<sup>[6]</sup>。然而,MLP由于自身较浅的网络结构,很难学习到图像的深度特征,并且随着神经元个数以及网络层数的增加,容易出现过拟合的现象。随着神经科学的发展,一种能够进行深度学习的神经网络,卷积神经网络(convolutional neural network)正发挥着越来越重要的作用,CNN正在与计算机视觉与模式识别相关领域取得令人瞩目的效果。由于遥感图像获取数据的特点,结合遥感数据源的CNN相关研究正在蓬勃发展,最初CNN在遥感图像方面的应用是在图像检索与场景分类上,Hu等<sup>[7]</sup>利用迁移学习的理念,从几种预先训练的大型网络中的各层中提取图像特征信息,用于遥感图像的场景分类,在几种公开数据集上的测试效果都显示的很好。葛芸等<sup>[8]</sup>通过在ImageNet数据集上预先训练4种卷积神经网络模型,将其运用于遥感图像检索,结果表明预先训练的卷积神经网络能有效的改进遥感图像检索性能。关于基于CNN的遥感图像目标检测也有较大的发展,姚相坤等<sup>[9]</sup>和黄洁等<sup>[10]</sup>利用CNN能够提取复杂图像特征的能力,将CNN用于舰船与飞机等识别度较高的物体检测,并取得较高的检测精度。Ren等<sup>[11]</sup>利用修正后的Faster R-CNN提高了小物体在遥感影像上检测的精度。然而,不同于CNN面对的传统图像场景分类或目标检测任务,当需要获取遥感影像上地物分布的格局与分布范围大小,做到精确测量时,在遥感影像分类中,要求做到对每一个像元进行精准分类。利用CNN对遥感影像的基于像元分类的研究引起了不少学者的兴趣,由于遥感数据源的形式多样,在对于中等分辨率的多光谱遥感影像,例如Landsat8, Sharma等<sup>[12]</sup>提出一种基于图像斑块大小为5的CNN系统,对Landsat8影像进行分类,取得较好的分类效果。付秀丽等<sup>[13]</sup>基于LandsatTM5数据源,以图像斑块大小为32作为输入构建CNN系统,对影像进行分类,最终实验精度表明,CNN分类精度明显高于与之对比的SVM算法。对于高分辨率卫星影像,例如worldview2, Zhang等<sup>[14]</sup>构建了一种MLP与CNN相结合的分类器对城市与郊区的两种不同类型的地物类型进行分类,分类效果得到很大的提高。曹林林等<sup>[15]</sup>基于quickbird数据源采用面向对象的分类方法,通过对比CNN和SVM两种分类算法对城区地区分类效果与精度,实验结果表

明 CNN 的分类精度要明显高于 SVM。陈广胜等<sup>[16]</sup>提出基于遗传算法和卷积网络分类方法,对高分辨率遥感影像进行基于光谱-空间-语义特征分类,实验结果表明,提出的分类方法对不同类别样本的分类精度普遍较好。对于高光谱的数据的研究中,由于高光谱数据的高维度性,降低数据的维度,选取合适的波段是处理高光谱遥感影像分类的关键, Yue 等<sup>[17]</sup>展示了一种结合主成分分析、深度卷积神经网络与逻辑回归的学习框架,对高光谱遥感影像进行分类,取得了较高的分类效果。李竺强等<sup>[18]</sup>提出一种适合高光谱分类的深度卷积神经网络模型,模型框架充分利用高光谱影像中像素空间领域信息与光谱信息,在测试数据集上达到很高的分类精度。Långkvist 等<sup>[19]</sup>探究了不同结构的 CNN 对遥感影像分类与分割的影响,对 CNN 的结构参数对分类影响有着全面的分析。

利用遥感影像进行植被分类的研究也在不断的发展,包括不同数据源与不同分类算法的比较<sup>[20]</sup>。Dalponte 等<sup>[21]</sup>通过将高密度激光雷达(light detection and ranging)数据分别和高光谱与多光谱数据进行融合运用于植被类型以及树种分类,表明了高密度 LiDAR 数据在植被分类上的潜力。多光谱影像由于相较于高光谱以及 LiDAR 数据在获取和处理方面容易,运用多光谱影像数据源植被分类是首先被探索的,但由于高空间分辨率而一般光谱分辨率较低,探索的方向侧重在挖掘其丰富的纹理信息,利用灰度共生矩阵(gray level co-occurrence)提取纹理特征用于分类是大多数学者研究的方向。从纹理特征的筛选<sup>[22]</sup>、纹理方向<sup>[23]</sup>以及移动窗口大小<sup>[24]</sup>等几个方向探索提高分类精度是研究的热点,但是利用 CNN 网络自生提取影像空间的高级空间特征用于分类的研究还较少,尚有发展的潜力。

本文主要研究以 MLP 为基准、比较不同图像斑块大小为输入的 CNN 之间在高分二号影像上林分类型分类效果的差异性,利用开源的机器学习框架 Tensorflow 搭建 CNN 与 MLP 模型,以高分二号 4 波段地物反射率作为样本,探究 CNN 模型在高分二号卫星遥感影像分类上的表现。具体探讨的问题是:(1)比较两种神经网络(CNN 和 MLP),在高分影像林分类型分类上的分类效果;(2)比较探讨不同图像斑块大小的 CNN 的分类效果,得出适合高分二号林分类型分类的 CNN。

## 1 数据源和研究区

本实验的研究区选取三明市将乐县国有林场周边的部分林分。实验区域林分内的树种含针叶林与

阔叶林。针叶林主要是各个龄级的杉木(*Cunninghamia lanceolata*)和马尾松(*Pinus massoniana*);阔叶林中包含木荷(*Schima superba*)和泡桐(*Paulownia*),还有其他的阔叶树散生在路边或沟谷里。根据实际的调查以及分类的要求,本文的林分类型分类包括灌木、马尾松林、其他阔叶林、裸地和杉木林。

遥感影像的数据源是高分二号卫星的全色与多光谱影像。成像时间是 2017 年 5 月 14 日。多光谱扫描仪获得的数据含 4 个光谱波段和一个全色波段。光谱波段包括蓝(450~520 nm)、绿(520~590 nm)、红(630~690 nm)和近红外(770~890 nm)。全色波段的空间分辨率为 1.0 m,光谱波段的分辨率为 4.0 m。在遥感图像的预处理阶段,首先,将高分二号影像的全色与多光谱影像利用自带的 RPC 信息和数字高程模型进行正射校正处理,然后将多光谱影像在 ENVI 中利用 FLAASH 模块进行大气校正处理,最后利用 Gran-Schmidt 将全色与多光谱影像进行融合获得 0.96 m 的高分多光谱融合影像。

在本实验中我们根据森林资源小班调查图和地面样地调查数据获取训练样本和测试样本,并用于接下来的模型的输入与精度的验证。森林资源小班调查图由当地的林业部门提供,样地调查数据更具小班分布图设置样地,对样地内树木进行树种鉴别,胸径,树高以及利用 GPS 定位。并在遥感图像进行标记,确保精确的匹配。通过在遥感影像的标记,我们在通过在遥感影像上选取感兴趣区(ROIs),将样本分为训练样本和测试样本,并且利用 Python 中 scikit-learn 将训练数据打乱并且按照 3:1 拆分为训练和验证数据(表 1),用于在训练模型的过程中判断是否过拟合和验证。

## 2 方法和原理

### 2.1 MLP 模型

MLP 是一种将输入数据以一种前馈方式映射到输出的网络,尽管在本篇文章中人工神经元不是感知器,我们还是称呼这种网络为 MLP。这种网络由相互连接的多层神经元组成,包括输入层、隐藏层和输出层(图 1)。人工神经元的类型取决于激活函数,感知器和 sigmoid 神经元都是激活函数,每个神经元的输出处理方式取决于激活函数,例如感知器的输出是 0 或 1。而 sigmoid 神经元的输出是就不仅仅是 0 或 1 这两个数,它能够输出 0 到 1 中的任意一个数,这一点有利于机器学习。通常,激活函数的神经元输出形式是:

$$f(x) = f(w_i x_i + b_i) \quad (1)$$

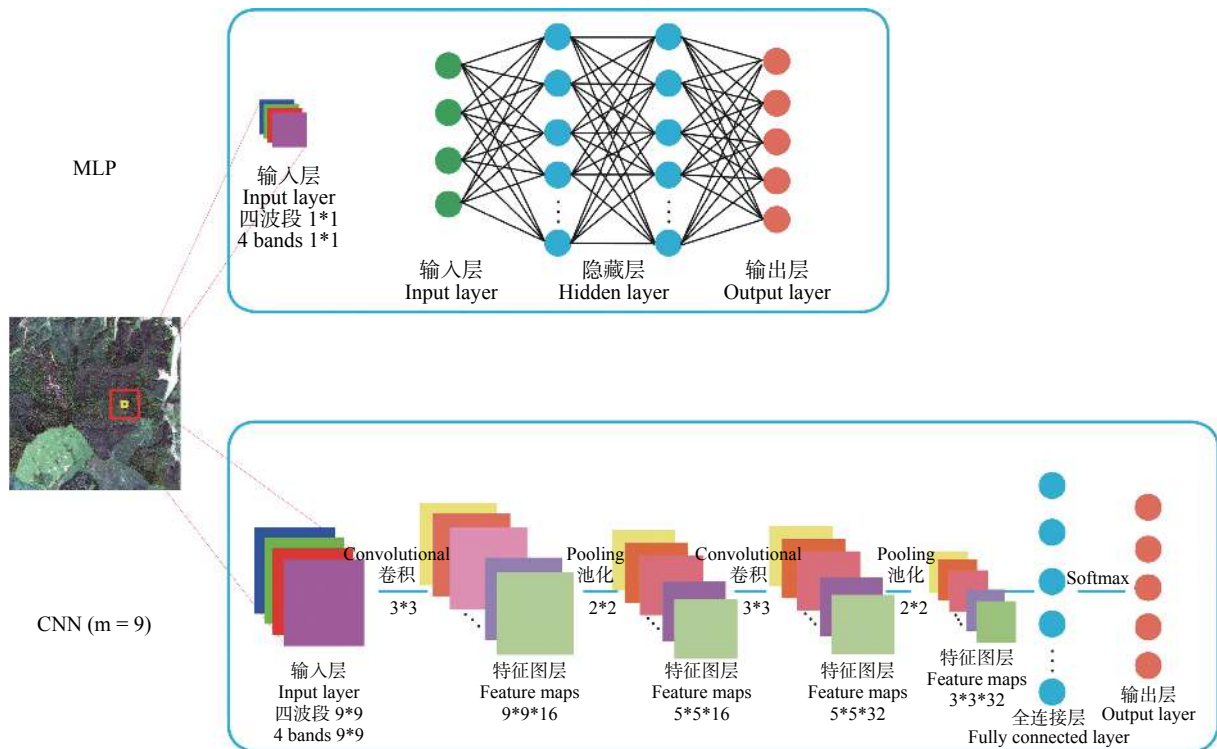
式中: $x_i$ ,  $w_i$  和  $b_i$  分别代表在  $i$  层中输入值、权重值和



表 1 研究区分类类别描述及其训练、验证和测试样本大小

Tab. 1 Description of class at study with training, validation and testing samples size

类别名称 Class	详细描述 Descriptions	训练样本 Train samples	验证样本 Validationsamples	测试样本 Testsamples
灌草 Shrub	林中草地, 生长大量灌草的伐后林地, 布满杂草的林道 Grass glade, nonforested land with large amount bushes and grasses, forest road with large shrubs	1 653	525	690
马尾松 Masson's pine	不同龄级的马尾松林, 主要是近熟林和成熟林 Masson's pine forest with different age class, mainly half-mature forest and mature forest	2 008	661	1 289
其他阔叶林 Other broadleaf	一些阔叶树, 主要是由木荷组成的林带, 散生的泡桐以及其他阔叶树种组成的混交林 Some broad-leaf trees, including schima used as fire belt, scattered paulownia and other mixed broadleaf forest	1 823	590	1 190
裸地 Bare soil	主要是几乎不含植被的农业用地和林业用地 Cropland and forest land with low percentage or no vegetation	1 653	575	347
杉木林 Chinese fir	不同龄级的杉木林, 主要是幼龄林、中龄林和成熟林 Chinese fir forest with different age class, including young forest, half-mature forest and mature forest	2 743	943	1 548

图 1 MLP 和 CNN( $m=9$ )的结构框架Fig. 1 Architecture of the CNN ( $m=9$ ) and MLP

偏差值,  $f$  代表着激活函数。

损失函数或代价函数是一种函数, 这种函数能够找到权重与偏差值使得从网络中输出的值与实际的标签值无限接近, 从而达到学习的目的。

$$C(w, b) = \frac{1}{2n} \sum_x y(x) - a^2 \quad (2)$$

式中:  $w$  代表网络中所有的权重值;  $b$  代表所有的偏差值;  $n$  输入网络中训练样本的数量;  $a$  是当  $x$  作为输入时, 网络的输出值;  $y(x)$  代表  $x$  作为对应的标签值。

我们称  $C$  为二次损失函数, 是一种在早期运用

最多的损失函数, 而另外一种损失函数的有效形式是交叉熵函数:

$$C = -\frac{1}{n} \sum_x [y \ln a + (1-y) \ln (1-a)] \quad (3)$$

式中:  $n$  代表训练数据的总数,  $x$  是输入值,  $y$  是对应的标签值,  $a$  是  $f(x)$ 。

## 2.2 CNN 模型

利用全连接层的神经网络例如 MLP 用于图像分类, 往往取得的效果不好, 是因为网络的框架没有考虑图像的空间结构信息, 因而提出一种适合图像

分类的新型神经网络, 卷积神经网络(CNN)这种网络能够快速的学习和训练, 卷积神经网络有3个基础的概念: 局部感受野, 权重共享和池化。

**局部感受野:**与MLP相比, 在隐藏层中的神经元不是直接和输入神经元进行一对一链接, 而是通过将输入神经元组成一个小的区域对接隐藏层中的神经元, 该区域称之为局部感受野, 感受野大小将影响获取图像空间信息的能力。

**共享权重和偏差值:**通常我们将从输入层输出到隐藏层的小区域称之为“特征图层”, 我们将“特征图”的权重称之为共享权重, “特征图层”的偏差值称之为共享偏差值。共享的权重与偏差通常被称之为“核”或“滤波器”(filter)。

$$a^1 = f(b + w * a^0) \quad (4)$$

式中:  $a^1$  代表一个“特征图层”的输出值,  $a^0$  代表输入的特征图层,  $*$ 代表卷积操作。

**池化层:**除了上述的卷积操作, CNN也包含池化操作, 池化操作一般在卷积操作之后, 池化操作的主要目的是简化从卷积操作输出的信息, 池化层将从卷积层中输出的特征图结果转化为更加浓缩的特征图层。例如, 在池化层中的每一个单元总结上一个图部分的神经元的信息, 一个最常见的池化操作是最大池化操作, 在最大池化操作当中, 池化单元总结上一个图部分区域的神经元的最大值最为最终的输出值, 池化的大小一般是  $2 \times 2$ , 也就是说将上一个图层中的4个神经元中的最大值将代表这  $2 \times 2$  的区域特征, 从而缩小简化了图的信息。

当经过卷积与池化操作后, 网络结构将进入全连接层, 这部分的信息传递与MLP类似。相比较于MLP, CNN框架的更加的复杂, 但变量的数量较少, 获得的信息较多, 所以容易训练与学习, 从而达到较高的分类效果。

### 2.3 精度评价指标

**总体精度(OA):**OA是指被正确分类的像元总数与总像元数的比值。

**Kappa系数(Kappa):**它是通过把所有地表真实分类中的像元总数乘以混淆矩阵对角线的和, 再减去某一类地表真实像元总数与该类中被分类像元总数之积对所有类别求和的结果, 再除以总像元数的平方减去某一类真实像元总数与该类中被分类像元总数之积对所有类别求和的结果所得到的。

**制图精度(PA):**PA是指某类别中被正确分类的像元数与该类别真实参考总数的比值。

**用户精度(UA):**UA是指某类别中被正确分类的像元数与被分为该类别像元总数的比值。

**Fscore:**Fscore是将制图精度与用户精度相结合的一个分类精度评价指标, Fscore越大精度越高, 其值在0到1之间, 公式如下:

$$Fscore = 2 * \frac{PA * UA}{PA + UA} \quad (5)$$

## 3 模型框架

### 3.1 模型的输入与输出

不同于普通的图像, 遥感图像数据格式独特, 因此需要设计一种合适的框架进行有效的学习。高分二号遥感影像在本质上是一个大小为  $X \times Y \times Z$  的多维数组。MLP和CNN均以4个波段的反射率作为模型的输入, 不同的是, MLP只利用一个像元的数据, 而构建的CNN则是以不同图像斑块大小作为输入。图像斑块大小  $m = 5, 7, 9, 11$ , 对应的输入数据为  $5 \times 5 \times 4, 7 \times 7 \times 4, 9 \times 9 \times 4$  和  $11 \times 11 \times 4$  多维数组的数据。MLP和CNN的输入值对应的标签值为斑块的中心像元的类别。两类模型的输出是5种分类的类别, 其形式是一个含各类别的概率(0到1)的一维数组。

### 3.2 模型的构造和超参数

图1展示了CNN( $m = 9$ )和MLP的总体框架。在CNN( $m = 9$ )网络系统中, 卷积层的层数设置为2层包含了卷积操作和池化操作, 选择2层能够更好地保持模型的复杂度和鲁棒性。在卷积操作中, filter的大小设置为  $3 \times 3$ , 为获取邻近像元的更多空间信息, 池化操作选择最大池化, 池化单元选择  $2 \times 2$ 。在卷积操作后, 卷积的结果输入到全连接层, 全连接层神经元个数设置为1000, 框架的最后一层为softmax层, 神经元的个数设置为5, 对应分类的5中类别。在整个框架中出去最后一层, 其他层的激活函数都选择的是修正线性单元函数(rectified linear unit, RELU)。相比较于sigmoid激活函数, RELU不会导致饱和, 学习不会下降。RELU已经被广泛的使用在各个神经网络当中。在选择代价函数的时候我们选择的是交叉熵函数而不是二次代价函数, 选择交叉熵函数有利于模型的更好的学习。优化器(optimizer)选择的是ADAM优化器, 学习率为0.0001, 权重的初始化方式选择截断分布。为实施以上的模型, 我们使用一款开源的机器学习框架Tensorflow和Python的GDAL、Numpy包用于搭建神经网络和遥感数据的读取与输出。图1中展示了MLP模型的框架, 从左至右, 绿色的神经元是输入层, 神经元个数为4, 网络中含2个隐藏层(颜色为蓝色的神经元), 每层含有100个神经元, 颜色为橘色的神经元为输出层, 神经元个数设置为5, 对应分

类的类别数。激活函数、权重初始化方式、代价函数以及优化器(这里学习率为 0.001)和 CNN 的配置相同。

### 3.3 模型的训练与验证

过拟合是神经网络训练过程中存在的最主要的问题。由于现在流行的一些神经网络大都含有较多的权重值与偏差值并且网络结构很深,模型过于复杂,在面对较小的数据集时,很容易产生过拟合的问题。为了更加有效的训练和提高模型的性能,我们要在训练的过程中监测过拟合是否发生,一旦过拟合发生我们就停止训练。在本文中,我们采用早停(Early Stopping)和 Dropout 两种方式防止过拟合。首先我们利用验证数据,在每一次训练迭代后进行验证精度,一旦我们发现验证精度和损失值在接下来的 100 次迭代中不在变化我们就停止训练,保存模型的框架与权重偏差值。在这里我们利用 tensorboard(一个在 Tensorflow 中用于可视化学习的模块)记录准确率与损失值的变化,其次为防止神经元数量过多,导致模型过于复杂而产生的过拟合现象,我们利用 Dropout 的方式:在每次迭代的种随机的丢弃全连接层的 50% 的神经元,剩余的神经元用

于训练,这样在每次迭代时就如同使用不同的结构,大量的实验证明,Dropout 是一种有效的防止过拟合的方式,另外防止过拟合的方式还有很多,例如正则化等。防止过拟合是神经网络训练的一项必要的措施。图 4 所展示的是 CNN( $m=9$ )模型的训练过程中准确率、损失值的变化曲线以及早停点的位置图。在迭代第 130 次时,损失值和准确率在接下来的 100 次中未有明显的变化,我们选择停止迭代,保存权重与偏差值固定模型用于接下来的影像分类。

## 4 结果与分析

我们利用测试数据与生成的分类图形成混淆矩阵来进行分类结果的准确率评估。评价的指标有制图精度(PA)、用户精度(UA)、总体精度(OA)、Kappa 系数和 Fscore。表 2~6 列出了 5 种神经网络的下的混淆矩阵,表 7 给出了 5 种神经网络分类系统的分类准确度对比。

从表 7 的数据我们可以看出:CNN( $m=9$ )分类精度达到最高,OA 和 Kappa 系数高达 94.78% 和 0.931 8,高于 MLP(83.87%, 0.790 4)、CNN( $m=5$ )

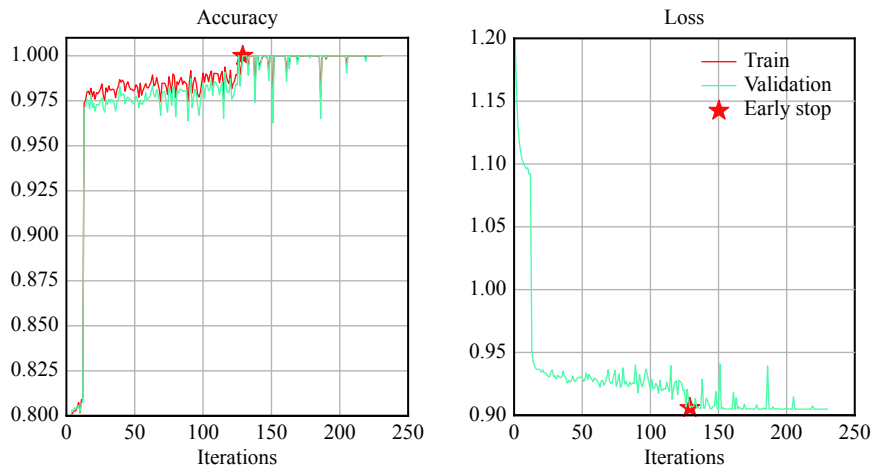


图 2 CNN( $m=9$ )在训练阶段准确率与损失值图

Fig. 2 Accuracy (left) and loss (right) of the CNN ( $m=9$ ) in stage of training

表 2 基于 MLP 生成的混淆矩阵

Tab. 2 Confusion matrix using the MLP

类别 Class	Ground Truths (pixels)				
	灌木 Shrub	马尾松林 Masson's pine	其他阔叶树林 Other broadleaf	裸地 Bare soil	杉木林 Chinese fir
灌木 Shrub	515	0	340	0	0
马尾松林 Masson's pine	0	1 200	0	0	130
其他阔叶树林 Other broadleaf	169	0	844	0	20
裸地 Bare soil	0	0	0	342	2
杉木林 Chinese fir	6	121	6	5	1 252
生产者精度 Producer's accuracy/%	74.64	90.84	70.92	98.56	89.17
用户精度 User's accuracy/%	60.23	90.23	81.70	99.42	90.07

表 3 基于 CNN 生成的混淆矩阵 ( $m = 5$ )Tab. 3 Confusion matrix using the CNN ( $m = 5$ )

类别 Class	Ground Truths (pixels)				
	灌草 Shrub	马尾松林 Masson's pine	其他阔叶树林 Other broadleaf	裸地 Bare soil	杉木林 Chinese fir
灌草 Shrub	<b>619</b>	0	300	0	0
马尾松林 Masson's pine	0	<b>1 190</b>	0	0	53
其他阔叶树林 Other broadleaf	67	0	<b>873</b>	0	1
裸地 Bare soil	0	0	0	<b>337</b>	0
杉木林 Chinese fir	4	131	17	10	<b>1 350</b>
生产者精度 Producer's accuracy/%	89.71	90.83	73.36	97.12	96.15
用户精度 User's accuracy/%	67.36	95.74	92.77	100	89.29

表 4 基于 CNN 生成的混淆矩阵 ( $m = 7$ )Tab. 4 Confusion matrix using the CNN ( $m = 7$ )

类别 Class	Ground Truths (pixels)				
	灌草 Shrub	马尾松林 Masson's pine	其他阔叶树林 Other broadleaf	裸地 Bare soil	杉木林 Chinese fir
灌草 Shrub	<b>620</b>	0	126	0	0
马尾松林 Masson's pine	0	<b>1 263</b>	0	0	9
其他阔叶树林 Other broadleaf	69	0	<b>1 064</b>	0	1
裸地 Bare soil	0	0	0	<b>325</b>	<b>0</b>
杉木林 Chinese fir	1	58	0	22	<b>1 647</b>
生产者精度 Producer's accuracy/%	89.86	95.61	89.41	93.66	96.65
用户精度 User's accuracy/%	83.11	97.45	92.76	100	94.37

表 5 基于 CNN 生成的混淆矩阵 ( $m = 9$ )Tab. 5 Confusion matrix using the CNN ( $m = 9$ )

类别 Class	Ground Truths (pixels)				
	灌草 Shrub	马尾松林 Masson's pine	其他阔叶树林 Other broadleaf	裸地 Bare soil	杉木林 Chinese fir
灌草 Shrub	<b>605</b>	0	59	0	0
马尾松林 Masson's pine	0	<b>1 288</b>	0	0	30
其他阔叶树林 Other broadleaf	78	11	<b>1 095</b>	0	19
裸地 Bare soil	0	0	0	<b>342</b>	0
杉木林 Chinese fir	7	13	36	5	<b>1 355</b>
生产者精度 Producer's accuracy/%	87.68	98.17	92.02	98.56	96.51
用户精度 User's accuracy/%	91.11	97.72	91.02	100	95.69

表 6 基于 CNN 生成的混淆矩阵

Tab. 6 Confusion matrix using the CNN ( $m = 11$ )

类别 Class	Ground Truths (pixels)				
	灌草 Shrub	马尾松林 Masson's pine	其他阔叶树林 Other broadleaf	裸地 Bare soil	杉木林 Chinese fir
灌草 Shrub	<b>531</b>	0	1	0	0
马尾松林 Masson's pine	0	<b>1 262</b>	0	0	52
其他阔叶树林 Other broadleaf	120	4	<b>1 137</b>	0	24
裸地 Bare soil	0	0	0	<b>293</b>	0
杉木林 Chinese fir	39	37	52	54	<b>1 328</b>
生产者精度 Producer's accuracy/%	76.96	96.85	95.55	84.44	94.59
用户精度 User's accuracy/%	99.81	96.04	88.48	100	87.95



表 7 不同分类系统分类精度对比利用每种类别的 Fscore、OA 和 Kappa

Tab. 7 Classification accuracy comparison amongst MLP, CNN ( $m = 5$ ), CNN ( $m = 7$ ), CNN ( $m = 9$ ) and CNN ( $m = 11$ ) for study area using per-class F score, overall accuracy and kappa coefficient. The bold font highlights the greatest classification accuracy per row.

类别 Class	Classifier				
	MLP	CNN ( $m = 5$ )	CNN ( $m = 7$ )	CNN ( $m = 9$ )	CNN ( $m = 11$ )
灌草 Shrub	0.666 7	0.769 4	0.863 5	<b>0.893 6</b>	0.869 1
马尾松林 Masson's pine	0.905 3	0.928 2	0.965 2	<b>0.979 4</b>	0.964 4
其他阔叶树林 Other broadleaf	0.759 3	0.819 3	0.910 6	0.915 2	<b>0.918 8</b>
裸地 Bare soil	0.989 9	0.985 4	0.967 3	<b>0.992 7</b>	0.915 6
杉木林 Chinese fir	0.896 2	0.925 9	0.955 0	<b>0.961 0</b>	0.911 5
总体精度 Overall accuracy/%	83.87	88.23	93.48	<b>94.78</b>	92.24
Kappa 系数 Kappa coefficient	0.790 4	0.847 1	0.914 9	<b>0.931 8</b>	0.897 9

(88.23%, 0.9471)、CNN( $m = 7$ )(93.48%, 0.914 9)和 CNN( $m = 11$ )(92.24%, 0.897 9)。可以看出 5 种分类系统的准确率排序为: MLP < CNN( $m = 5$ ) < CNN( $m = 11$ ) < CNN( $m = 7$ ) < CNN( $m = 9$ )。从表 7 中我们可以看到 CNN( $m = 9$ )不仅在 OA 和 Kappa 系数上表现的高于其他分类系统,在对每种类别的分类准确率上也较高。为了综合比较每种类别的分类准确度,我们利用 Fscore 来评价, Fscore 结合了 PA 和 UA,能更好地反映分类的准确度。从表中的数据中,我们可以看出, CNN( $m = 9$ )在每种类别上的分类精度都高于 CNN( $m = 5$ )和 MLP,其中在灌草这一类别上, CNN( $m = 9$ )比 CNN( $m = 5$ )和 MLP 要分别高出 0.124 2 和 0.226 9。当我们将 CNN( $m = 9$ )和 CNN( $m = 11$ )相比较的时候,我们发现,除其他阔叶树林这一类别外,在其他所有类别上, CNN( $m = 9$ )的分类准确率都要大于 CNN( $m = 11$ ),并且 CNN( $m = 11$ )在其他阔叶树林这一类别上的分类准确率仅仅比 CNN( $m = 9$ )稍微大一点(0.003 6),所以我们

可以认为 CNN( $m = 9$ )在每种类别上的分类精度是要优于 CNN( $m = 11$ )。当与 CNN( $m = 7$ )对比的时候我们发现,虽然在每种类别的分类精度上, CNN( $m = 9$ )与 CNN( $m = 7$ )的准确度非常的接近,但我们也能看出在每种类别上都有微量的提升。当我们将 CNN( $m = 5$ ,  $m = 7$  和  $m = 11$ )与 MLP 准确度对比时,我们发现,除去裸地这一类别外, CNN 在每种类别上的分类准确度是要高于 MLP 的,这里对于裸地这一类别,我们看见 MLP 的准确度是要明显高于 CNN,其中比 CNN( $m = 11$ )要高 0.074 3。在对比除 CNN( $m = 9$ )以外的 CNN 的分类准确度上,我们可以看出其中 CNN( $m = 7$ )在灌草、马尾松林、裸地和杉木林 4 种类别上都要比 CNN( $m = 11$ )和 CNN( $m = 5$ )要高。

图 3 展示了研究区的两个子图(A 和 B)的分类图,首先,从 MLP 的分类结果图中我们可以清楚的看到,在 A 和 B 上都显示出不同程度的噪点,在 A、B 中,我们看到杉木林类别被错误分类为马尾松林

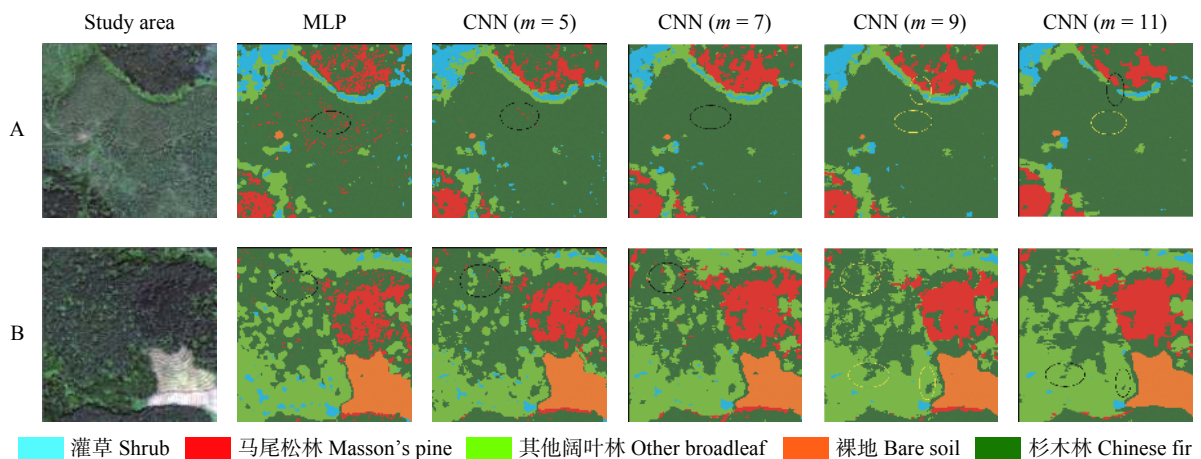


图 3 两幅研究区子图在不同分类系统下的分类结果

Fig. 3 Two image subsets (A and B) in study area with classification results. Columns from left to right represent the original images, the MLP classification and CNN ( $m = 5, 7, 9, 11$ ) classification correspondingly.



类别,在B中一些其他阔叶林林类别也被误分类为灌草类别。但是我们也可以看出,MLP在鉴别裸地这一类别时,表现的很好。接着,我们将CNN的分类图与MLP相互比较,我们能够看出,CNN分类图中图像“噪点”在减少,平滑的效果在明显的增加。当我们对比CNN( $m=5,7$ )和MLP时,虽然平滑效果在增加,但“噪点”还是存在,尽管CNN( $m=7$ )的效果比CNN( $m=5$ )要好。并且在结合CNN( $m=9$ )和CNN( $m=11$ ),我们发现,随着图像斑块大小的增加,出现了另外一种问题,即由于过度平滑而产生的类别间的界限变得不确定,一些小的或线状的类别被擦除合并到一些大的类别当中去。例如我们在观察CNN( $m=11$ )的分类图中,我们在A中,发现由其他阔叶林组成的林带被切断并入到杉木类别当中,在B中一些小部分的杉木类别被并入到其他阔叶林类别当中,类别间的边界变得不确定性。在我们观察CNN( $m=9$ )的分类图时,我们发现“噪点”和边界不确定明显的问题得到很大的减弱,在CNN( $m=11$ )中被误分的林带和小部分的杉木也得到正确的分类,在MLP和CNN( $m=5,7$ )中出现的不同程度的“噪点”也到明显的减少。

## 5 讨论与结论

MLP作为一种传统的神经网络模型,依靠的是单个像元的光谱信息并没有挖掘高分辨率遥感影像丰富的空间信息特征,因此MLP高分影像树种识别上未取得所期望的效果,OA(83.97%)和Kappa(0.7904),远低于CNN( $m=9$ )(94.78%,0.9318)。对于大面积的存在的杉木林类别以及其他阔叶林类别,虽然能够鉴别正确,但由于椒盐现象而产生的“噪点”,MLP在两种类别的分类精度(0.8962,0.7593)上都低于CNN( $m=9$ )(0.9610,0.9152)。

CNN,作为一种现今热门的深度学习神经网络,能够从高分辨卫星影像上挖掘高级的空间特征,同时,特有的卷积与池化操作能够在减少变量的维度的同时,最大可能的总结图像的特征信息。由于CNN的这些优势,在本文中,CNN( $m=5,7,9,11$ )的分类准确率都要比MLP要高,分类效果要好。然而事物都有两面性,CNN也不能例外,随着图像斑块大小的增加,尽管椒盐现象的影响在逐渐的减弱,但由于过度平滑而产生的边界不确定性的问题随之而来。为避免两种分类问题的干扰从而达到较高的分类准确率,选择一个适合的图像斑块大小作为输入的CNN是关键,实验的结果也表明,当我们将图像的斑块大小设置为9时,分类的准确度和分类图的分类效果达到最好。也就是说图像斑块大小在

CNN遥感影像分类中至关重要,斑块大小过小,会导致挖掘的空间信息不够,出现如同MLP一样的椒盐现象,当斑块的大小过大,会出现过度平滑的效果,导致对面积小的或线性的类别被忽略,从而降低分类的准确度。尺度的选择是根据实验所获得的精度数据以及分类图的可视化得出的,但我们结合数据源与研究的分类对象也能得到启发,数据源的特征,例如空间分辨率的大小,以及分类对象特点是影响图像斑块( $m$ )的重要因素。

在本文当中,我们提出一种为高分二号卫星遥感影像林分类型分类定制的CNN系统,该CNN系统以图像斑块大小为9输入,能够充分的挖掘高分影像的空间和光谱信息,通过实验的对比MLP以及其他斑块大小输入的CNN,CNN( $m=9$ )在总体精度与各类别的分类精度上都取得了较大的提升。在这其中,仍有许多值得继续研究深入的。一方面,卷积神经网络的框架以及超参数的选择,一直是热门的研究方向,调整参数与框架一直是神经网络的难点。根据数据特征,选择合适的框架与参数显得尤为重要。另一方面,在遥感影像数据源方面,随着遥感理论与传感器技术的发展,数据源的类型多样,了解不同数据源的优势,利用数据源之间的互补作用,结合CNN分类会发挥巨大的提升。再者,结合各种分类算法特点,利用CNN能够提取高级的空间特征,使用其他分类算法进行最后决策分类,也是一个值得升入研究的方向。

## 参 考 文 献

- [1] 樊文锋,李鸿洲,温奇,等. 高分二号卫星影像正射纠正精度分析[J]. 测绘通报, 2016(9): 63-66.  
Fan W F, Li H Z, Wen Q, et al. Research on accuracy of digital Ortho Rectification of GF-2 image[J]. Bulletin of Surveying and Mapping, 2016(9): 63-66.
- [2] 陈业培,孙开敏,白婷,等. 高分二号影像融合方法质量评价[J]. 测绘科学, 2017(11): 35-40.  
Chen Y P, Sun K M, Bai T, et al. Quality assessment on image fusion methods for GF-2 data[J]. Science of Surveying and Mapping, 2017(11): 35-40.
- [3] Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: a review[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66(3): 247-259.
- [4] Rodriguez-Galiano V F, Chica-Olmo M, Abarca-Hernandez F, et al. Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture[J]. Remote Sensing of Environment, 2012, 121: 93-107.
- [5] Mas J F, Flores J J. The application of artificial neural networks to the analysis of remotely sensed data[J]. International Journal of Remote Sensing, 2008, 29(3): 617-663.
- [6] Kumar U, Kerle N, Punia M, et al. Mining land cover information

- using multilayer perceptron and decision tree from MODIS data[J]. *Journal of the Indian Society of Remote Sensing*, 2010, 38(4): 592–603.
- [7] Hu F, Xia G, Hu J, 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.
- [8] 葛芸, 江顺亮, 叶茂茂, 等. 基于 ImageNet 预训练卷积神经网络的遥感图像检索 [J]. 武汉大学学报(信息科学版), 2018, 43(1): 67–73.
- Ge Y, Jiang S L, Ye F M, et al. Remote sensing image retrieval using pre-trained convolutional neural networks based on ImageNet[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(1): 67–73.
- [9] 姚相坤, 万里红, 霍宏, 等. 基于多结构卷积神经网络的高分遥感影像飞机目标检测 [J]. 计算机工程, 2017(1): 259–267.
- Yao X K, Wan L H, Huo H, et al. Airplane object detection in high resolution remote sensing imagery based on multi-structure convolutional neural network[J]. *Computer Engineering*, 2017(1): 259–267.
- [10] 黄洁, 姜志国, 张浩鹏, 等. 基于卷积神经网络的遥感图像舰船目标检测 [J]. 北京航空航天大学学报, 2017(9): 1841–1848.
- Huang J, Jiang Z G, Zhang H P, et al. Ship object detection in remote sensing images using convolutional neural networks[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2017(9): 1841–1848.
- [11] Ren Y, Zhu C, Xiao S. Small object detection in optical remote sensing images via modified faster R-CNN[J]. *Applied Sciences*, 2018, 8(5): 813.
- [12] Sharma A, Liu X, Yang X, et al. A patch-based convolutional neural network for remote sensing image classification[J]. *Neural Networks*, 2017, 95: 19–28.
- [13] 付秀丽, 黎玲萍, 毛克彪, 等. 基于卷积神经网络模型的遥感图像分类 [J]. 高技术通讯, 2017(3): 203–212.
- Fu X L, Li L P, Mao K B, et al. Remote sensing image classification based on CNN model[J]. *Chinese High Technology Letters*, 2017(3): 203–212.
- [14] Zhang C, Pan X, Li H, et al. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 140: 133–144.
- [15] 曹林林, 李海涛, 韩颜顺, 等. 卷积神经网络在高分遥感影像分类中的应用 [J]. 测绘科学, 2016(9): 170–175.
- Cao L L, Li H T, Han Y S, et al. Application of convolutional neural networks in classification of high resolution remote sensing imagery[J]. *Science of Surveying and Mapping*, 2016(9): 170–175.
- [16] 陈广胜, 陈守玉, 景维鹏, 等. 基于遗传算法与全卷积网络的遥感影像像素级分类方法 [J]. 小型微型计算机系统, 2018(7): 1590–1595.
- Chen G S, Chen S Y, Jing W P, et al. Remote sensing image pixel-level classification with genetic algorithm and fully convolutional network[J]. *Journal of Chinese Computer Systems*, 2018(7): 1590–1595.
- [17] Yue J, Mao S, Li M. A deep learning framework for hyperspectral image classification using spatial pyramid pooling[J]. *Remote Sensing Letters*, 2016, 7(9): 875–884.
- [18] 李竺强, 朱瑞飞, 高放, 等. 三维卷积神经网络模型联合条件随机场优化的光谱遥感影像分类 [J]. 光学学报, 2018(8): 404–413.
- Li Z Q, Zhu G F, Gao F, et al. Hyperspectral remote sensing image classification based on three-dimensional convolution neural network combined with conditional random field optimization[J]. *Acta Optica Sinica*, 2018(8): 404–413.
- [19] Långkvist M, Kiselev A, Alirezaie M, et al. Classification and segmentation of satellite orthoimagery using convolutional neural networks[J]. *Remote Sensing*, 2016, 8(4): 329.
- [20] Fassnacht F E, Latifi H, Stereńczak K, et al. Review of studies on tree species classification from remotely sensed data[J]. *Remote Sensing of Environment*, 2016, 186: 64–87.
- [21] Dalponte M, Bruzzone L, Gianelle D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data[J]. *Remote Sensing of Environment*, 2012, 123: 258–270.
- [22] 王妮, 彭世揆, 李明诗. 基于树种分类的高分辨率遥感数据纹理特征分析 [J]. 浙江农林大学学报, 2012(2): 210–217.
- Wang N, Peng S K, Li M S. High-resolution remote sensing of texture images for tree species classification[J]. *Journal of Zhejiang A & F University*, 2012(2): 210–217.
- [23] Fu F, Wan X J, Wang J, et al. Tree species and age groups classification based on GF-2 image[J]. *Remote Sensing for Land and Resources*, 2019, 31(2): 118–124.
- [24] 李丹, 柯樱海, 宫辉力, 等. 基于高分辨率遥感影像的城市典型乔木树种分类研究 [J]. 地理与地理信息科学, 2016(1): 84–89.
- Li D, Ke Y H, Gong H L, et al. Urban tree species classification with machine learning classifier using WorldView-2 imagery[J]. *Geography and Geo-Information Science*, 2016(1): 84–89.

(责任编辑 冯秀兰

责任编辑委)