

基于卷积神经网络的高分遥感影像单木树种分类

欧阳光^{1,2}, 荆林海^{1*}, 阎世杰¹, 李慧¹, 唐韵玮¹, 谭炳香³¹中国科学院空天信息创新研究院数字地球重点实验室, 北京 100094;²中国科学院大学电子电气与通信工程学院, 北京 100049;³中国林业科学研究院资源信息研究所, 北京 100091

摘要 树种调查一直面临着成本高、效率低、精度不高等问题。利用遥感手段能大大提高树种类型调查的工作效率、节省成本;卷积神经网络(CNN)虽然已经在自然图像分类领域取得了许多突破,但是较少有人将 CNN 模型用于单木树种分类。基于上述考虑,搭建出 CNN 模型,并与高分遥感影像相结合,进行单木树种分类。在利用高分影像半自动化构建单木树种遥感影像样本集过程中,采用了影像冠层切片(CSI)圈定、人工标注、数据增强等方法;同时为了训练单木树种遥感影像样本集,对 5 个 CNN 模型进行针对性改写。通过对比分析发现:LeNet5_relu 和 AlexNet_mini 都未取得最佳分类效果;GoogLeNet_mini56、ResNet_mini56 和 DenseNet_BC_mini56 分别对不同的树种具有最佳分类效果;DenseNet_BC_mini56 总体精度最高(94.14%),Kappa 系数最高(0.90),是总体最佳分类模型。该研究证明了 CNN 在单木树种分类中的有效性,能为森林资源调查提供重要的解决方案。

关键词 图像处理;单木树种分类;卷积神经网络;高分遥感影像;深度学习

中图分类号 TP79

文献标志码 A

doi: 10.3788/LOP202158.0228002

Classification of Individual Tree Species in High-Resolution Remote Sensing Imagery Based on Convolution Neural Network

Ouyang Guang^{1,2}, Jing Linhai^{1*}, Yan Shijie¹, Li Hui¹, Tang Yunwei¹, Tan Bingxiang³¹Key Laboratory of Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;²School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Science, Beijing 100049, China;³Institute of Forest Resource Information Techniques CAF, Beijing 100091, China

Abstract Tree species investigation has been faced with problems such as high cost, low efficiency, and low precision. The use of remote sense can greatly increase the work efficiency of tree species investigation and save cost. Although convolutional neural network (CNN) has made many breakthroughs in natural image classification area, few people have used CNN model to carry out individual tree species classification. Based on the above considerations, this paper builds CNN models, and integrates them with high-resolution remote sensing imagery to classify individual tree species. In the course of semi-automatically constructing the sample set of remote sensing imagery of individual tree species with high-resolution imagery, the crown slices from imagery (CSI) delineation, manual annotation, and data augmentation are used. Meanwhile, in order to train the sample set of remote sensing imagery of individual tree species, five CNN models are adapted. Through comparative analysis, it is found that LeNet5_relu and AlexNet_mini cannot achieve the best classification effect. GoogLeNet_mini56, ResNet_mini56, and DenseNet_BC_mini56 have the best classification effect for different species respectively. DenseNet_BC_mini56

收稿日期: 2020-06-12; 修回日期: 2020-06-23; 录用日期: 2020-07-03

基金项目: 中国科学院空天信息创新研究院重点部署项目(Y951150Z2F)、中国科学院战略性先导科技专项(A类)子课题(XDA19030501)、新疆维吾尔自治区重大专项课题(2018A03004)、国家自然科学基金面上项目(41972108)

* E-mail: jinglh@radi.ac.cn

has the highest overall accuracy (94.14%) and the highest Kappa coefficient (0.90), making it the best classification model from all aspects. The research proves the effectiveness of CNN in the classification of individual tree species, which can provide a critical solution for forest resource investigation.

Key words image processing; individual tree species classification; convolution neural network; high-resolution remote sensing imagery; deep learning

OCIS codes 100.2000; 100.4996; 280.4788

1 引言

森林生态系统是地球上最大的陆地生态系统,也是全球生物圈中重要的一环,在改善生态环境、维持生态平衡、保护人类生存发展基本环境等方面发挥着不可替代的作用^[1]。森林资源管理是林业部门的主要工作之一,其中树种类型调查对于森林资源管理和监测具有十分重要的意义^[2]。传统的树种类型调查,主要采用实地调查及林相图辅助等方法,不仅消耗大量的人力和物力,而且更新困难、准确率较低,无法满足林业部门以及用户对于森林遥感信息精细提取的要求^[3-4]。

高分辨率遥感影像包含了丰富的纹理、形状、位置等信息,不仅能显示出较大的地貌特征,还能清晰地反映森林中单棵树木的细节,这些细节为单木树冠的高精度提取、单木参数的估计以及单木尺度的树种分类等提供了可能,利用高分辨率遥感影像对森林区域进行树种类型调查已经成为相关研究领域的热点^[5-7]。

传统的树种分类主要采用了人工设计特征结合机器学习分类器的方法,诸如 K 最邻近(KNN)^[8]、决策树(DT)^[9]、随机森林(RF)^[10]、最大似然法(MLC)^[11]、支持向量机(SVM)^[12]、人工神经网络(ANN)^[13]等。Franklin 等^[14]使用 RF 对无人飞行器采集的多光谱影像中 4 类田间树种进行面向对象分类,总体准确性约为 78%;Immitzer 等^[15]使用 RF 对 WorldView2 影像中 10 类森林树种进行了分类,总体准确性约为 82%;Li 等^[16]利用 SVM 和 RF 的方法对双时域 WorldView2/3 影像中城市树种进行了分类,结果表明,SVM 比 RF 更加适合样本分布不均衡的城市树种分类;Wang 等^[17]基于 Pleiades-1 高分辨率遥感数据,采用 DT、SVM 及 RF 对红树林进行树种识别,描绘出 80% 以上的红树林;谢珠利^[18]使用了 KNN、DT、RF、MLC、SVM、ANN 等 6 个分类器,对国产 ZY-3 卫星影像中 6 类人工林树种进行了提取,结果表明,不同的分类算法有各自的优缺点,可能某种算法对某类树种有较好的分类效果;李永亮等^[19]将高光谱特征参数作为输

入矢量,森林树种类别作为输出矢量,建立反向传播(BP)神经网络模型,对三种树木进行监督分类,分类综合精度达 93.3%;余超^[20]以 WorldView2 影像作为数据源,分别采用 DT、SVM 和 ANN 这 3 种分类器对森林公园主要树种进行分类,结果表明,DT 分类精度最高,SVM 次之,ANN 最差。以上方法中分类特征的构建依赖于主观设置的经验和参数值,分类器仅能从原始图像中提取到浅层次的地物特征用于分类,而对于面向对象的深层次的地物特征表达不够充分,分类效果有限^[21]。

随着深度学习的发展,卷积神经网络(CNN)近年来广泛应用于图像分类领域^[22-23]。深度学习 CNN 模型利用多层神经网络结构充分学习图像中包含的特征信息,自动提取出更抽象、更深层的信息用于图像分类,因此,CNN 往往具有比机器学习更好的分类效果。已经有少部分学者开始探索将深度学习用于遥感影像的树种识别与分类,均取得了不错的结果^[24-25]。Wegner 等^[26]通过车载激光雷达(LiDAR)收集城市的树木类别,并使用 CNN 模型进行分类实验,实验结果表明,基于 CNN 模型的对象检测器和分类器可以显著提高树种的分类精度,从而替代人工树种调查;Zou 等^[27]利用 LiDAR 采集树木的三维空间信息,运用深度信念网络(DBN)进行分类,对树种的分类精度高达 95.6%;Onishi 等^[28]利用无人机高分影像、数字高程模型(DEM)及坡度模型提取树冠,并进行基于对象的 CNN 模型分类,对 7 类树种达到了 89% 的总体精度;滕文秀等^[29]结合面向对象和深度迁移学习方法对 QuickBird 高分影像中的森林公园进行树种分类,达到了 78.98% 的总体精度,远远高于 RF 和 SVM 的分类结果。已有深度学习单木树种分类多利用机载 LiDAR 数据,这类数据源的获取成本高、处理复杂,且覆盖范围有限,对于精细的单木树种分类应用来说,采用高分影像更为妥当。深度学习模型虽然展现出优于机器学习分类器的单木树种分类效果,但是对单木树种样本集的数量有一定的要求。因此,遥感影像的单木树种分类在数据源选择、样本集构建、精度提升等方面还存在改进的地方。

本文针对森林地区单木树种分类存在的问题,将 CNN 模型与高分影像相结合,构建典型森林地区的单木树种遥感影像样本集,探索最优的 CNN 模型,实现典型森林地区单木树种的最佳分类方案,提升树种类型调查的自动化水平,为森林资源管理及监测提供技术支持。

2 研究区与实验数据

2.1 研究区概况

研究区选在安徽黄山风景区,图 1 展示了研究区所在的位置,坐标范围为 $118^{\circ}9'16''\text{E} \sim 118^{\circ}11'24''\text{E}$, $30^{\circ}7'8''\text{N} \sim 30^{\circ}10'37''\text{N}$,占地面积约为

21.6 km^2 。由于景区地形呈山峰凸起、沟长谷深的特征,一般峰顶为黄山松纯林,峰体为黄山松(或马尾松)与阔叶树混交林,峰底(谷、沟)为阔叶树纯林,人工乔木林主要分布于景区外围地带。

2.2 实验数据

遥感数据选取为一景 WorldView3 影像,成像时间为 2019 年 3 月 10 日。该景影像全色波段空间分辨率为 0.3 m ,8 个多光谱波段空间分辨率为 1.2 m 。为了最大程度保留光谱和纹理信息,对影像进行了基于雾霾和波段比值(HR)^[30]的融合处理,将融合后 0.3 m 的 8 波段影像用于接下来的分类。

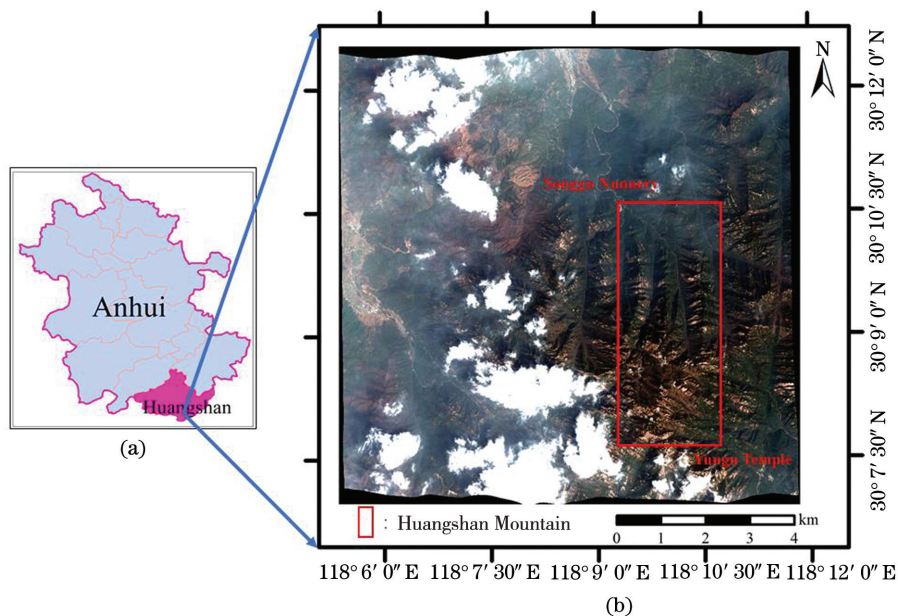


图 1 研究区所在位置。(a)黄山市在安徽省的方位;(b)WorldView3 真彩色示意图,方框为黄山风景区所在位置
Fig. 1 Location of the research area. (a) Huangshan City, Anhui Province; (b) true color schematic of WorldView3, the box indicates the location of Huangshan Mountain

3 样本集构建与划分

3.1 单木树种遥感影像样本集的构建

利用 CNN 模型对高分影像进行单木树种分类,首先需要大量单木树种遥感影像样本用于训练以及测试模型,因此,本文在研究区内建立相应单木树种遥感影像样本集。单木树种遥感影像样本集需要从整幅遥感影像中裁剪出仅包含单棵树冠的影像并确定其类别,而遥感影像覆盖范围广,地物类型比较复杂,对遥感影像中树冠的描绘需要采取一定的方法,本文采取的影像冠层切片(CSD)^[31]算法是一种专门进行树冠圈定的高效算法,能对遥感影像进行多尺度分割,自动圈定影像

中的树冠。

本文单木树种遥感影像样本集的构建分为:1)在研究区实施野外实地调查,采集树种类别样本;2)对研究区遥感影像进行 CSI 树冠圈定处理,得到树冠圈定的矢量数据;3)将矢量数据叠加到影像上,结合遥感影像解译结果,标注树冠的类别;4)标注类别后的树冠取最小外接矩形后输出,得到单木树种遥感影像图;5)根据树种类别进行归类,构建包含研究区 5 类主要树种的单木树种遥感影像样本集。如图 2 所示。

黄山风景区范围较大,地形复杂,影像云层覆盖较多,实地考察后,将野外采样区域选在观景路线两侧。2019 年 12 月底,课题组成员在采样区域使用

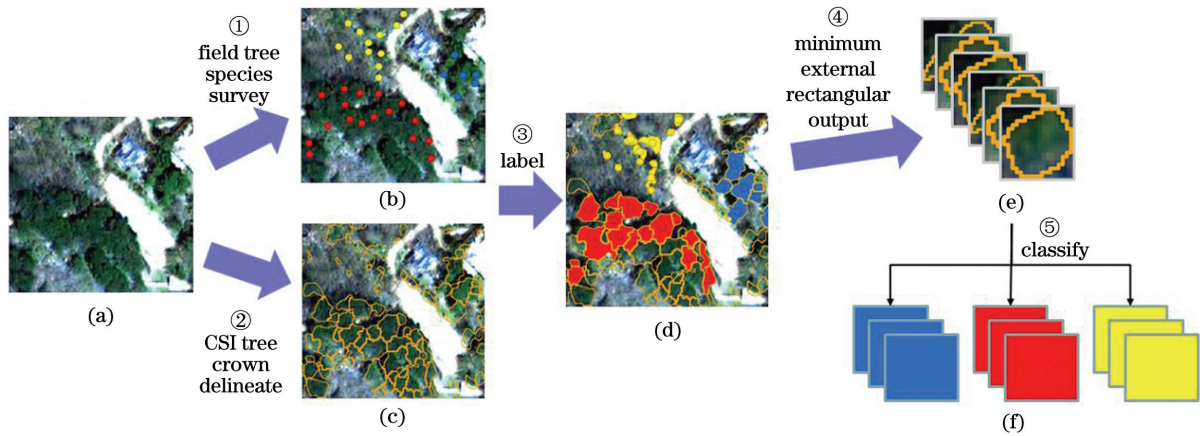


图 2 单木树种遥感影像样本集的构建步骤。(a)研究区遥感影像;(b)树种分布图;(c)树冠圈定图;(d)树冠类别标注图;(e)单木树种遥感影像图;(f)单木树种遥感影像样本集






Fig. 2 Construction steps of sample set of remote sensing imagery of individual tree species. (a) Remote sensing imagery of research area; (b) distribution diagram of tree species; (c) delineation diagram of tree crown; (d) labeling diagram of tree crown category; (e) remote sensing imagery of individual tree species; (f) sample set of remote sensing imagery of individual tree species

全球定位系统(GPS)手持机采取了一定的样本点,作者根据野外采样结果对树种进行了合并,并对影像进行了类别标注后归类,结果如表1所示。单木

树种遥感影像样本集共分为 5 类,其中毛竹(Ph. h) 112 个、常绿乔木(E. a)44 个、杉木(C. l)139 个、黄山松(Pi. h)2001 个、落叶乔木(D. a)617 个。

表 1 单木树种遥感影像样本集类别标注结果

Table 1 Classification labeling result of sample set of remote sensing imagery of individual tree species

Tree species	Species merged	Shorthand	Field sampling points	Classification labeled sample set	Remote sensing imagery
<i>Phyllostachys heterocycla</i>	<i>Phyllostachys heterocycla</i>	Ph. h	18	112	
<i>Daphniphyllum oldhamii</i> / <i>Daphniphyllum macropodum</i> / <i>Osmanthus fragrans</i> / <i>Manglietia sp.</i> / <i>Ilex chinensis</i> / <i>Camellia japonica</i> / <i>Rhododendron sp.</i>	Evergreen arbor	E. a	113	444	
<i>Tsuga chinensis</i> / <i>Abies fabri</i> / <i>Taxus sp.</i>	<i>Cunninghamia lanceolata</i>	C. l	67	139	
<i>Pinus hwangshanensis</i>	<i>Pinus hwangshanensis</i>	Pi. h	245	2001	
<i>Cyclobalanopsis glauca</i> / <i>Emmenopterys henryi</i> / <i>Tilia japonica</i> / <i>Castanea seguinii</i> / <i>Acer sp.</i> / <i>Sorbus sp.</i>	Deciduous arbor	D. a	327	617	
Total	—	—	703	3313	—

3.2 样本集划分

在深度学习中,通常将样本集随机划分为训练样本集、验证样本集、测试样本集。首先使用训练样本集训练模型,然后用验证样本集验证模型,根据验证误差不断调整模型,选出其中的最优模型,最后用测试样本集评估模型。本文的样本集划分配比为训练样本集:验证样本集:测试样本集=3:1:1。

CNN 模型通常具有几万甚至数百万的参数,需要有足够比例的样本参与训练,通过野外采样后标注的样本数很难达到相应的规模,因此,还需要对训练样本集和验证样本集进行数据增强。本文对原始的单木树种遥感影像进行 90° 旋转、180° 旋转、270° 旋转、上下翻转、左右翻转等操作后,样本集的数量扩大为原来的 6 倍。表 2 为经过数据增强前后样本集的划分情况。

表 2 数据增强前后的样本集划分结果

Table 2 Results of sample set division before and after data augmentation

Tree species	Training sample set		Validation sample set		Test sample set
	Before	After	Before	After	
Ph. h	66	396	23	138	23
E. a	266	1596	89	534	89
C. l	83	498	28	168	28
Pi. h	1199	7194	401	2406	401
D. a	369	2214	124	744	124
Total	1983	11898	665	3990	665

4 模型与方法

传统 CNN 模型的输入一般为单通道灰度图像集或三通道彩色图像集,尺度为固定的 32 或 224;而单木树种遥感影像样本集的波段数为 8,尺度为 6~60 不等。因此,单木树种遥感影像样本集输入模型前需要提前缩放到统一尺度下。为了照顾不同尺度的样本,方便 CNN 模型的改写,本文将这一尺度固定为 32。基于上述考虑,本文参考 CNN 模型发展阶段中出现的几个重要模型,在保留模型核心思想的前提下,通过改写模型参数,搭建出适合单木树种遥感影像样本集多波段、小尺度特点的模型。

最早用于识别手写数字(MNIST)图像集,是 CNN 模型的开山之作,确定了 CNN 模型的基本架构。

本文模型的搭建策略为:1)将输入层由适合灰度图像的 32×32 改为适合多光谱影像的 32×32×8;2)激活函数由 tanh 改为 ReLU。为方便区分,本文中搭建后的模型命名为 LeNet5_relu,尾标“_relu”表示激活函数使用的是 ReLU。

4.1 LeNet5 模型

LeNet5 模型是 1998 年由 LeCun 等^[32]提出的,

LeNet5_relu 模型共 8 层,包含输入层、3 个卷积层、2 个池化层、1 个全连接层和分类层。其中第 2、4、6 层为卷积层,分别包含 6、16、120 个大小为 5×5,步长均为 1 的卷积滤波器,激活函数均为 ReLU;第 3、5 层为池化层,都采用 2×2 的平均池化,步长均为 2;第 7 层为全连接层,包含 84 个输出神经元节点;分类层采用 Softmax 分类器分出 5 类。如表 3 所示。

表 3 LeNet5_relu 模型参数

Table 3 LeNet5_relu model parameter

Layer	Output size	Parameter
Input	32×32×8	—
Convolutional C1	28×28×6	Kernel 5×5, filter 6, stride 1, ReLU
Pooling S1	14×14×6	Average_pooling 2×2, stride 2
Convolutional C2	10×10×16	Kernel 5×5, filter 16, stride 1, ReLU
Pooling S2	5×5×16	Average_pooling 2×2, stride 2
Convolutional C3	1×1×120	Kernel 5×5, filter 120, stride 1, ReLU
Fully-connected F1	84	Node 84, FC, ReLU
Classification	5	Node 5, FC, Softmax

4.2 AlexNet 模型

AlexNet^[33]模型是由 2012 年 ImageNet 大规模视觉识别挑战 (ILSVRC) 分类任务冠军获得者 Hinton 和他的学生 Krizhevsky 所设计。模型引入 ReLU 加速训练,通过重叠池化提取特征,加入局部响应归一化(LRN)层增加模型泛化能力,使用数据增强和 Dropout 降低过拟合。AlexNet 模型最初用于 $224 \times 224 \times 3$ 的 ImageNet 图像集 的识别,模型中网络的卷积核和输出均较大,因此,需要缩小网络的规模。

本文模型搭建的策略为:1)将输入层由 $227 \times 227 \times 3$ 改为 $32 \times 32 \times 8$;2)将 C1 层卷积核由 11×11 减小为 7×7 ,防止过大大局部感受野导致的欠拟合;

表 4 AlexNet_mini 模型参数

Table 4 AlexNet_mini model parameter

Layer	Output size	Parameter
Input	$32 \times 32 \times 8$	—
Convolutional C1	$32 \times 32 \times 12$	Kernel 7×7 , filter 12, stride 1, ReLU
Pooling S1	$15 \times 15 \times 12$	Average_pooling 3×3 , stride 2
Convolutional C2	$15 \times 15 \times 36$	Kernel 5×5 , filter 36, stride 1, ReLU
Pooling S2	$7 \times 7 \times 36$	Average_pooling 3×3 , stride 2
Convolutional C3	$7 \times 7 \times 54$	Kernel 3×3 , filter 54, stride 1, ReLU
Convolutional C4	$7 \times 7 \times 54$	Kernel 3×3 , filter 54, stride 1, ReLU
Convolutional C5	$3 \times 3 \times 36$	Kernel 3×3 , filter 36, stride 1, ReLU
Pooling S3	$3 \times 3 \times 36$	Average_pooling 3×3 , stride 2
Fully-connected F1	320	Node 320, FC, ReLU
Fully-connected F2	100	Node 100, FC, ReLU
Classification	5	Node 5, FC, Softmax

4.3 GoogLeNet 模型

GoogLeNet^[34]模型是由 2014 年 ILSVRC 分类任务冠军获得者 Szegedy 等设计。模型设计了 Inception 模块,可在保持计算成本的同时增加网络的深度和宽度。GoogLeNet 模型也是用于 $224 \times 224 \times 3$ 的 ImageNet 图像集 的识别,同样需要减少卷积核的数量和输出。除此以外,GoogLeNet 模型采用了 6 次下采样,尺度大小为 32 的图像经过 5 次下采样已经缩减为 1 pixel,因此,还需要减少下采样的次数。

本文模型搭建的策略为:1)将输入层由 $224 \times 224 \times 3$ 改为 $32 \times 32 \times 8$;2)在输入层和 Inception 层之间只保留 C1 层,且 C1 层步长由 2 改为 1,减少下采样次数;3)C1 层卷积滤波器数量由 64 减少为 12,Inception 模块内所有卷积滤波器数量减为原来

3)所有卷积滤波器数量减少 $1/7 \sim 1/6$,全连接层输出由 4096 减少到 320 和 100,防止过拟合;4)去掉 Dropout 层。搭建后的模型命名为 AlexNet_mini,尾标“_mini”表示模型参数规模被大大缩小。

AlexNet_mini 模型共 12 层,包含输入层、5 个卷积层、3 个池化层、2 个全连接层和分类层。其中第 2、4、6、7、8 层为卷积层 C1~C5,分别包含 12 个大小为 7×7 ,36 个大小为 5×5 ,54、54、36 个大小为 3×3 ,步长均为 1 的卷积滤波器,激活函数均为 ReLU;第 3、5、9 层为池化层 S1~S3,采用 3×3 的平均池化,步长均为 2;第 10、11 层为全连接层 F1 和 F2,分别包含 320 和 100 个输出神经元节点;分类层采用 Softmax 分类器分出 5 类。如表 4 所示。

的 $1/8$,防止过拟合。搭建后的模型命名为 GoogLeNet_mini56,尾标“_mini56”表示模型不仅参数规模被大大缩小,而且卷积层和全连接层加起来共 56 层。

GoogLeNet_mini56 模型包含输入层、1 个卷积层、9 个 Inception V1 模块、3 个池化层和分类层。其中第 2 层为卷积层 C1,包含 12 个大小为 7×7 ,步长为 1 的卷积滤波器,激活函数为 ReLU;第 1、2 Inception V1 模块分别为 1a 和 1b,第 3~7 Inception V1 模块分别为 2a~2e,第 8、9 Inception V1 模块分别为 3a 和 3b;池化层 S1 和 S2 分别在 Inception V1 模块 1b 和 2e 后,采用 3×3 的最大池化,步长均为 2;池化层 S3 在 Inception V1 模块 3b 后,采用 7×7 的平均池化,步长为 1;分类层采用 Softmax 分类器分出 5 类。如表 5 所示。

表 5 GoogLeNet_mini56 模型参数
Table 5 GoogLeNet_mini56 model parameter

Layer	Output size	Parameter
Input	$32 \times 32 \times 8$	—
Convolutional C1	$32 \times 32 \times 12$	Kernel 7×7 , filter 12, stride 1
Inception V1 block (1a)	$32 \times 32 \times 32$	—
Inception V1 block (1b)	$32 \times 32 \times 60$	—
Pooling S1	$15 \times 15 \times 60$	Max_pooling 3×3 , stride 2
Inception V1 block (2a)	$15 \times 15 \times 64$	—
Inception V1 block (2b)	$15 \times 15 \times 64$	—
Inception V1 block (2c)	$15 \times 15 \times 64$	—
Inception V1 block (2d)	$15 \times 15 \times 66$	—
Inception V1 block (2e)	$15 \times 15 \times 104$	—
Pooling S2	$7 \times 7 \times 104$	Max_pooling 3×3 , stride 2
Inception V1 block (3a)	$7 \times 7 \times 104$	—
Inception V1 block (3b)	$7 \times 7 \times 128$	—
Pooling S3	$1 \times 1 \times 128$	Average_pooling 7×7 , stride 1
Classification	5	Node 5, FC, Softmax

对于 Inception V1 模块,它是由四部分并联而成。其中第一部分为大小 1×1 的卷积滤波器;第二部分使用大小 1×1 的卷积滤波器构成的瓶颈层,作为 3×3 卷积滤波器之前的降维;第三部分同样使用

了瓶颈层,作为 5×5 卷积滤波器之前的降维;第四部分同样使用了瓶颈层,作为大小 3×3 、步长为 1 的最大池化层之后的降维。Inception V1 模块中卷积滤波器的具体数量记录在表 6 中。

表 6 Inception V1 模块参数
Table 6 Inception V1 block parameter

Layer	Inception V1 block						
	I	II	III	IV			
	Convolutional kernel 1×1	Bottleneck kernel 1×1	Convolutional kernel 3×3	Bottleneck kernel 1×1	Convolutional kernel 5×5	Pooling 3×3	Bottleneck kernel 1×1
1a	Filter 8	Filter 12	Filter 16	Filter 2	Filter 4	Stride 1	Filter 4
1b	Filter 16	Filter 16	Filter 24	Filter 4	Filter 12	Stride 1	Filter 8
2a	Filter 24	Filter 12	Filter 26	Filter 2	Filter 6	Stride 1	Filter 8
2b	Filter 20	Filter 14	Filter 28	Filter 3	Filter 8	Stride 1	Filter 8
2c	Filter 16	Filter 16	Filter 32	Filter 3	Filter 8	Stride 1	Filter 8
2d	Filter 14	Filter 18	Filter 36	Filter 4	Filter 8	Stride 1	Filter 8
2e	Filter 32	Filter 20	Filter 40	Filter 4	Filter 16	Stride 1	Filter 16
3a	Filter 32	Filter 20	Filter 40	Filter 4	Filter 16	Stride 1	Filter 16
3b	Filter 48	Filter 24	Filter 48	Filter 6	Filter 16	Stride 1	Filter 16

4.4 ResNet 模型

ResNet^[35] 模型是由 2015 年 ILSVRC 分类任务冠军获得者 He 等设计。模型设计了 Residual 模

块,解决了深层网络的退化问题。ResNet 模型也适用于 $224 \times 224 \times 3$ 的 ImageNet 图像集的识别,同样需要减少卷积核的数量和输出。为更好说明不同

网络模型结构的效果,网络层数都设为 56 层;除此以外,56 层的 ResNet 模型会大大增加参数计算量,还需要在 Residual 模块内引入瓶颈层。

本文模型搭建的策略为:1)将输入层由 $224 \times 224 \times 3$ 改为 $32 \times 32 \times 8$;2)C1 层卷积滤波器数量由 64 减少为 12,Residual 模块内所有卷积滤波器数量减为原来的 1/8,防止过拟合;3)采用瓶颈层处理的 Residual 模块,模块重复参数设为 (3,4,8,3)。搭建后的模型命名为 ResNet_mini56,尾标“_mini56”与 GoogLeNet_mini56 的尾标同义,不再

赘述。

ResNet_mini56 模型包含输入层、1 个卷积层、1 个池化层、4 个 Residual 模块和分类层。其中第 2 层为卷积层 C1,包含 12 个大小为 7×7 、步长为 2 的卷积滤波器,激活函数为 ReLU;接下来为 Residual 模块(1)~(5);池化层 S1 在 Residual 模块(1)的第一层,采用 3×3 的最大池化,步长为 2;分类层先使用 384 个输出神经元节点的全局平均池化替代全连接层,最后采用 Softmax 分类器分出 5 类。如表 7 所示。

表 7 ResNet_mini56 模型参数

Table 7 ResNet_mini56 model parameter

Layer	Output size	Parameter	Residual blockfilter		
			I	II	III
			Bottleneck kernel 1×1	Convolutional kernel 3×3	Bottleneck kernel 1×1
Input	$32 \times 32 \times 8$	-		-	
Convolutional C1	$16 \times 16 \times 12$	Kernel 7×7 , filter 12, stride 2		-	
	Pooling S1	$8 \times 8 \times 12$		Max_pooling 3×3 , stride 2	
Residual block (1)	Bottleneck	$8 \times 8 \times 12$	12	-	-
	Convolutional	$8 \times 8 \times 12$	-	12	-
	Bottleneck	$8 \times 8 \times 48$	-	-	48
Residual block (2)	Bottleneck	$4 \times 4 \times 24$	24	-	-
	Convolutional	$4 \times 4 \times 24$	-	24	-
	Bottleneck	$4 \times 4 \times 96$	-	-	96
Residual block (3)	Bottleneck	$2 \times 2 \times 48$	48	-	-
	Convolutional	$2 \times 2 \times 48$	-	48	-
	Bottleneck	$2 \times 2 \times 192$	-	-	192
Residual block (4)	Bottleneck	$1 \times 1 \times 96$	96	-	-
	Convolutional	$1 \times 1 \times 96$	-	96	-
	Bottleneck	$1 \times 1 \times 384$	-	-	384
Classification	384	Global_average_pooling		-	
	5	Node 5, FC, Softmax			

对于 Residual 模块,除了采用 shortcut 的连接方式将模块输入与输出进行了恒等映射以外,还在模块内大小 3×3 的卷积滤波器之前使用瓶颈层进行了降维,并在之后通过瓶颈层恢复高维。Residual 模块中卷积滤波器的具体数量也记录在表 7 中。

4.5 DenseNet 模型

DenseNet 模型是由 2017 年 CVPR 最佳论文作者 Huang 等^[36]设计。模型设计了 Dense 模块和压

缩层,Dense 模块通过特征在通道上的密集连接实现特征复用,进一步解决了深层网络梯度消失的问题,压缩层通过降低特征图大小减少参数计算量。除此以外,同样可以在 Dense 模块内引入瓶颈层减少参数计算量。包含瓶颈层的 Dense 模块和压缩系数小于 1 的压缩层的组合模型称为 DenseNet_BC 模型。任意层的 DenseNet_BC 模型需满足(网络层数 - Dense 模块数 - 1)/(2 × Dense 模块数) = 整

数。DenseNet 模型既在 ImageNet 图像集进行过测试,也在 $32 \times 32 \times 3$ 的 CIFAR-10 图像集和 SVHN 图像集进行了测试。

本文模型搭建的策略为:1)将输入层改为适合多光谱影像的 $32 \times 32 \times 8$;2)C1 层卷积滤波器数量由 16 减少为 12,Dense 模块内每层卷积滤波器的增长数由 12 减少为 6,防止过拟合;3)采用瓶颈层处理的 Dense 模块,模块重复参数设为 5;4)压缩层系数设为 0.5。搭建后的模型命名为 DenseNet_BC_

mini56。

DenseNet_BC_mini56 模型包含输入层、1 个卷积层、5 个 Dense 模块、4 个压缩层和分类层。其中第 2 层为卷积层 C1,包含 12 个大小为 3×3 ,步长为 2 的卷积滤波器,激活函数为 ReLU;接下来为 Dense 模块(1)~(5)与压缩层(1)~(4)交替叠加;分类层先使用 99 个输出神经元节点的全局平均池化替代全连接层,最后采用 Softmax 分类器分出 5 类。如表 8 所示。

表 8 DenseNet_BC_mini56 模型参数

Table 8 DenseNet_BC_mini56 model parameter

Layer	Output size	Parameter	Dense block filter	
			I	II
			Bottleneck kernel 1×1	Convolutional kernel 1×1
Input	$32 \times 32 \times 8$	-	-	-
Convolutional C1	$32 \times 32 \times 12$	Kernel 3×3 , filter 12, stride 2	-	-
Dense block (1)	$32 \times 32 \times 42$	$\times 5$	24	-
			-	6
Compression (1)	$32 \times 32 \times 21$	Kernel 1×1	-	-
	$16 \times 16 \times 21$	Average_pooling 2×2 , stride 2	-	-
Dense block (2)	$16 \times 16 \times 51$	$\times 5$	24	-
			-	6
Compression (2)	$16 \times 16 \times 36$	Kernel 1×1	-	-
	$8 \times 8 \times 36$	Average_pooling 2×2 , stride 2	-	-
Dense block (3)	$8 \times 8 \times 66$	$\times 5$	24	-
			-	6
Compression (3)	$8 \times 8 \times 51$	Kernel 1×1	-	-
	$4 \times 4 \times 51$	Average_pooling 2×2 , stride 2	-	-
Dense block (4)	$4 \times 4 \times 81$	$\times 5$	24	-
			-	6
Compression (4)	$4 \times 4 \times 66$	Kernel 1×1	-	-
	$2 \times 2 \times 66$	Average_pooling 2×2 , stride 2	-	-
Dense block (5)	$2 \times 2 \times 99$	$\times 5$	24	-
			-	6
Classification	99	Global_average_pooling	-	-
	5	Node 5, FC, Softmax	-	-

对于 Dense 模块,除了采用密集连接方式将前面所有层和后面层建立关系以外,还在模块内大小 3×3 的卷积滤波器之前使用瓶颈层进行了降维。

Dense 模块的输出通道数由输入 Dense 模块的通道数和每层卷积滤波器的增长数共同决定,以 Dense 模块(2)为例,输入 Dense 模块(2)的通道数为 21,

每层卷积滤波器增长数为 6, 经过 5 层的生长, Dense 模块的输出通道数变为 $21+6\times 5=51$ 。压缩层的输出通道数由前一 Dense 模块的输入、输出通道数和压缩系数共同决定, 以压缩层(2)为例, 输入、输出 Dense 模块(2)的通道数分别为 21、51, 二者相加乘以压缩系数 0.5, 压缩层的输出通道数变为 $(21+51)/2=36$ 。

4.6 深度学习环境

本文采用前端 keras2.2.4 + 后端 TensorFlow 1.14.0 模式搭建了深度学习框架, 编程语言为

Python3.6.8, 操作系统为 Windows10, 显卡为 NVIDIA GTX 1060, 并通过 CUDA10.0 实现显卡的并行计算。

5 实验结果与分析

5.1 模型参数

模型参数的数量一定程度上反映了模型的复杂程度, 也决定了计算量。本文模型的参数摘要如表 9 所示。

表 9 CNN 模型参数

Table 9 CNN model parameter

Model name	Total parameter	Trainable parameter	Non-trainable parameter	Network layer
LeNet5_relu	62331	62331	0	5
AlexNet_mini	213537	213537	0	8
GoogLeNet_mini56	97251	97227	24	56
ResNet_mini56	934025	924401	9624	56
DenseNet_BC_mini56	82979	78839	4140	56

从表 9 可以看出: 1) LeNet5_relu 是最为简单的 5 层网络模型, 也是参数最少的模型; 2) ResNet_mini56 是最为复杂的 56 层网络模型, 也是参数最多的模型; 3) GoogLeNet_mini56 和 DenseNet_BC_mini56 虽然有 56 层网络, 但通过各自特殊的结构设计, 大大减少了参数量, 甚至连 8 层网络模型 AlexNet_mini 的 1/2 都不到; 4) ResNet_mini56 作为同样的 56 层网络模型, 但是参数量达到了 GoogLeNet_mini56 和 DenseNet_BC_mini56 的 10 倍左右, 模型训练的结果有可能会有过拟合的现象。

5.2 训练精度与验证精度

利用 3.2 节中划分好的训练样本集对模型进行

训练, 训练过程中同时使用验证样本集对训练好的模型进行精度验证。经过迭代训练, LeNet5_relu 模型在第 50 个周期收敛, 训练精度为 95.16%, 验证精度为 91.87%; AlexNet_mini 模型在第 41 个周期收敛, 训练精度为 95.23%, 验证精度为 92.77%; GoogLeNet_mini56 模型在第 71 个周期收敛, 训练精度为 95.12%, 验证精度为 94.59%; ResNet_mini56 模型在第 38 个周期收敛, 训练精度为 98.25%, 验证精度为 94.49%; DenseNet_BC_mini56 模型在第 42 个周期收敛, 训练精度为 98.87%, 验证精度为 97.21%。图 3 以柱形图形式表示了各模型收敛时的训练精度、验证精度与网络层数。

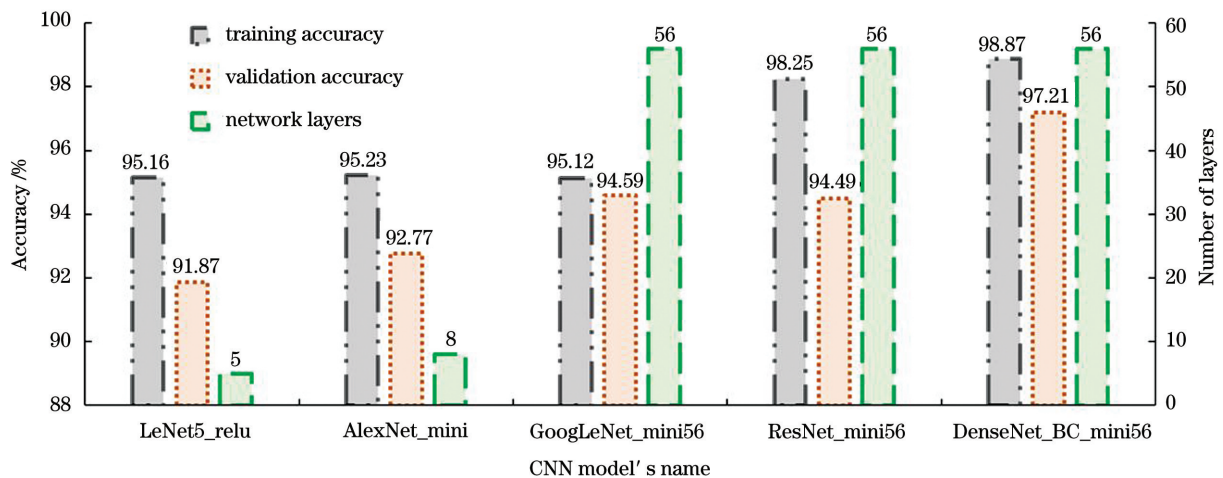


图 3 CNN 模型收敛时的训练精度、验证精度与网络层数柱形图

Fig. 3 Histogram of training accuracy, validation accuracy, and network layers when CNN model converges

通过对比模型的训练精度,可以看出:1)相比 LeNet5_relu 模型, AlexNet_mini 模型网络层次加深 3 层,训练精度略微提升;2)相比 LeNet5_relu 模型和 AlexNet_mini 模型, GoogleLeNet_mini56 模型虽然通过 Inception 模块将网络层次大大加深到 56 层,但是训练精度反倒下降,这是因为深层次网络出现了退化问题;3)相比 GoogleLeNet_mini56 模型, ResNet_mini56 具有相同的网络层次,训练精度大幅提升 3.13%,这是因为 Residual 模块解决了网络退化问题;4)相比 ResNet_mini56 模型, DenseNet_BC_mini56 同样具有相同的网络层次,训练精度进一步提升,达到最高 98.87%,这是因为通过 Dense 模块进一步解决了深层网络梯度消失的问题。

通过对比模型的验证精度,可以看出:1)前三个模型中, GoogleNet_mini56 模型最优, AlexNet_

mini 模型次之, LeNet5_relu 模型最差,说明模型网络层次越深,模型的验证效果越好;2)相比 GoogleLeNet_mini56 模型,虽然 ResNet_mini56 模型的训练精度远高于 GoogleLeNet_mini56 模型 3.13%,但是 ResNet_mini56 模型验证精度反而低 0.1%,这是因为 ResNet_mini56 模型参数过多,导致训练结果过拟合;3)相比其他所有模型, DenseNet_BC_mini56 模型的验证精度达到最高 97.21%,说明 DenseNet_BC_mini56 模型的验证效果最佳。

5.3 分类精度评价

为了更好地说明模型对不同树种进行分类预测的优劣势,本文使用测试样本集计算了各个模型的混淆矩阵,并根据混淆矩阵计算了生产者精度、用户精度、总体精度和 Kappa 系数等指标进行评价,如表 10 所示。

表 10 CNN 模型分类精度评价指标

Table 10 Classification accuracy evaluation index of CNN model

Model name	Evaluation index	Tree species				
		Ph. h	E. a	C. l	Pi. h	D. a
LeNet5_relu	Producer accuracy / %	86.96	67.42	75.00	98.00	87.90
	User accuracy / %	90.91	76.92	87.50	93.57	90.08
	Overall accuracy / %	90.68				
	Kappa coefficient	0.84				
AlexNet_mini	Producer accuracy / %	86.96	71.91	64.29	97.76	92.74
	User accuracy / %	90.91	79.01	78.26	94.00	94.26
	Overall accuracy / %	91.58				
	Kappa coefficient	0.85				
GoogLeNet_mini56	Producer accuracy / %	95.65	74.16	75.00	97.76	97.58
	User accuracy / %	100.00	84.62	95.45	94.92	93.08
	Overall accuracy / %	93.53				
	Kappa coefficient	0.89				
ResNet_mini56	Producer accuracy / %	95.65	76.40	78.57	97.51	92.74
	User accuracy / %	100.00	80.95	100.00	94.90	92.00
	Overall accuracy / %	92.93				
	Kappa coefficient	0.88				
DenseNet_BC_mini56	Producer accuracy / %	95.65	75.28	85.71	98.25	95.97
	User accuracy / %	100.00	87.01	100.00	94.71	94.44
	Overall accuracy / %	94.14				
	Kappa coefficient	0.90				

从表 10 可以看出:1)对任一树种类型来说, LeNet5_relu 和 AlexNet_mini 作为浅层网络模型都未取得最高的分类精度指标,分类效果都不是最佳,最佳分类效果在 56 层网络中产生;2)对毛竹来说, GoogLeNet_mini56、ResNet_mini56 和 DenseNet_BC_mini56 的生产者精度和用户精度最高且相同,都具有最佳的分类效果;3)对常绿乔木来说,ResNet_mini56 的生产者精度最高,具有最小漏分误差, DenseNet_BC_mini56 的用户精度最高,具有最小错分误差;4)对杉木来说,ResNet_mini56 的用户精度最高,具有最小的错分误差, DenseNet_BC_mini56 的生产者精度和用户精度都最高,具有最佳的分类效果;5)对黄山松来说, DenseNet_BC_mini56 的生产者精度最高,具有最小漏分误差, GoogLeNet_mini56 的用户精度最高,具有最小错分误差;6)对落叶乔木来说, GoogLeNet_mini56 的生产者精度最高,具有最小漏分误差, DenseNet_BC_mini56 的用户精度最高,具有最小错分误差;7)对所有模型来说, DenseNet_BC_mini56 的总体精度 94.14% 和 Kappa 系数 0.90 最高,因此, DenseNet_BC_mini56 是总体最佳分类模型。

5.4 黄山风景区树种分类图

基于 DenseNet_BC_mini56 模型的预测结果,本文对黄山风景区的树种进行了制图,如图 4 所示。

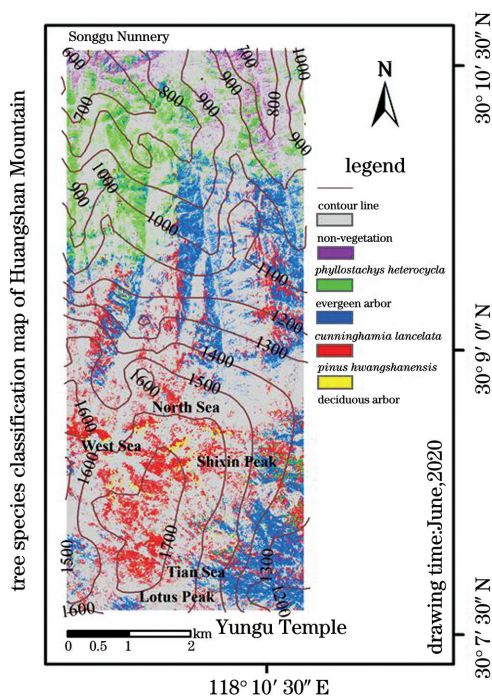


图 4 黄山风景区树种分类图

Fig. 4 Classification diagram of tree species of Huangshan Mountain

从图 4 中可以看出,黄山风景区树种呈现明显的海拔垂直分布特点,具体表现为:1)海拔 800 m 以下,主要为毛竹纯林、毛竹与常绿乔木混交林;2)海拔 800~1000 m,主要为杉木与常绿乔木混交林;3)海拔 1000~1200 m,主要为杉木与常绿乔木混交林、杉木与黄山松混交林,海拔 1000 m 开始出现黄山松;4)海拔 1200~1500 m,主要为黄山松与杉木混交林、黄山松与常绿乔木混交林,混交程度进一步加大;5)海拔 1500 m 以上,主要为黄山松纯林、黄山松与落叶乔木混交林。

6 结 论

本文为满足 CNN 模型对大量样本集的需求,采用 CSI 树冠圈定、人工标注、数据增强等方法半自动化构建出单木树种遥感影像样本集;针对样本集多波段、小尺度的特点,在原有 CNN 模型核心思想的基础上,通过制定针对性改写策略,搭建出 5 个 CNN 模型用于单木树种分类。经过多次迭代训练后再进行预测, CNN 模型实现了黄山风景区单木树种高精度分类。实验结果表明, DenseNet_BC_mini56 的训练精度为 98.87%, 验证精度为 97.21%, 总体精度为 94.14%, Kappa 系数为 0.90, 具有最佳的训练效果和总体分类效果。

本文研究目的是进行单木树种分类,而单木树种包含两方面的含义,一是尺度上的单棵树,二是种类上的单种树。除了 CNN 模型本身的深度特征提取优势外,本文的研究方法获得高精度单木树种分类有两个原因:1)采用了基于对象的分类。相比基于像素的分类,基于对象的分类往往可以获得更好的分类效果^[21]。本文通过 CSI 算法分割树冠,获得了树冠对象,但是该方法对阔叶树的分割效果一般。因此,改进分割方法可以在尺度上获得更高的精准度,更加精准地描绘单棵树、统计树种数量。2)进行了树种合并。从表 1 可以看出,除了毛竹和黄山松以外,常绿乔木、杉木以及落叶乔木都是对多类树种的合并,该策略考虑了 CNN 模型对样本集中单个类别的数量要求,但是会模糊树种信息。因此,在野外树种调查的过程中准确区分每一类树种,可以减少树种合并的类目,提高最终树种分类的精度。

另外,本文对影像进行 CSI 树冠圈定并标注类别,实际上是对影像进行语义分割的过程;全卷积神经网络(FCN)^[37]通过输入整幅原始影像,输出整幅语义分割的影像,实现端到端的分类效果。在计算机设备允许的情况下,可考虑设计出一种端到端的

FCN 模型,进一步提高应用效率。

参 考 文 献

- [1] Zhao T Q, Ouyang Z Y, Zheng H, et al. Forest ecosystem services and their valuation in China [J]. *Journal of Natural Resources*, 2004, 19(4): 480-491.
赵同谦, 欧阳志云, 郑华, 等. 中国森林生态系统服务功能及其价值评价 [J]. *自然资源学报*, 2004, 19(4): 480-491.
- [2] Song Q F. Study on impact of forest resource dynamic change on forest ecological function in recent 40 years in china [D]. Beijing: Chinese Academy of Forestry, 2015.
宋庆丰. 中国近 40 年森林资源变迁动态对生态功能的影响研究 [D]. 北京: 中国林业科学研究院, 2015.
- [3] Peng W. Research on the importance of forestry investigation planning and design in forestry production and construction [J]. *Journal of Green Science and Technology*, 2018(19): 141-142.
彭伟. 林业调查规划设计在林业生产建设中的重要性探究 [J]. *绿色科技*, 2018(19): 141-142.
- [4] Tang J H, Chen Y J. Analysis of the importance of forestry survey planning and design in forestry production and construction [J]. *Xiandai Horticulture*, 2015(20): 227-228.
唐景华, 陈育娟. 林业调查规划设计在林业生产建设中的重要性分析 [J]. *现代园艺*, 2015(20): 227-228.
- [5] Deng G. Research on individual tree identification and crown segmentation algorithm in high spatial resolution remote sensing imagery [D]. Beijing: Chinese Academy of Forestry, 2009.
邓广. 高空间分辨率遥感影像单株立木识别与树冠分割算法研究 [D]. 北京: 中国林业科学研究院, 2009.
- [6] Madonsela S, Cho M A, Mathieu R, et al. Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2017, 58: 65-73.
- [7] Majid I A, Latif Z A, Adnan N A. Tree species classification using worldview-3 data [C] // 2016 7th IEEE Control and System Graduate Research Colloquium (ICSGRC), August 8-8, 2016, Shah Alam, Malaysia. New York: IEEE Press, 2016: 73-76.
- [8] Blanzieri E, Melgani F. Nearest neighbor classification of remote sensing images with the maximal margin principle [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(6): 1804-1811.
- [9] Friedl M A, Brodley C E. Decision tree classification of land cover from remotely sensed data [J]. *Remote Sensing of Environment*, 1997, 61(3): 399-409.
- [10] Gislason P O, Benediktsson J A, Sveinsson J R. Random forests for land cover classification [J]. *Pattern Recognition Letters*, 2006, 27(4): 294-300.
- [11] Settle J J, Briggs S A. Fast maximum likelihood classification of remotely-sensed imagery [J]. *International Journal of Remote Sensing*, 1987, 8(5): 723-734.
- [12] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778-1790.
- [13] Civco D L. Artificial neural networks for land-cover classification and mapping [J]. *International Journal of Geographical Information Systems*, 1993, 7(2): 173-186.
- [14] Franklin S E, Ahmed O S. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data [J]. *International Journal of Remote Sensing*, 2018, 39(15/16): 5236-5245.
- [15] Immitzer M, Atzberger C, Koukal T. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data [J]. *Remote Sensing*, 2012, 4(9): 2661-2693.
- [16] Li D, Ke Y H, Gong H L, et al. Object-based urban tree species classification using Bi-temporal WorldView-2 and WorldView-3 images [J]. *Remote Sensing*, 2015, 7(12): 16917-16937.
- [17] Wang D Z, Wan B, Qiu P H, et al. Artificial mangrove species mapping using pléiades-1: an evaluation of pixel-based and object-based classifications with selected machine learning algorithms [J]. *Remote Sensing*, 2018, 10(2): 294-315.
- [18] Xie Z L. Plantation tree species classification with multi-source high resolution remotely sensed data [D]. Hangzhou: Zhejiang A & F University, 2019.
谢珠利. 基于多源高分辨率遥感数据的人工林树种分类研究 [D]. 杭州: 浙江农林大学, 2019.
- [19] Li Y L, Lin H, Sun H, et al. Study on forest species classification based on BP neural network [J]. *Journal of Central South University of Forestry & Technology*, 2010, 30(11): 43-46.
李永亮, 林辉, 孙华, 等. 基于 BP 神经网络的森林树种分类研究 [J]. *中南林业科技大学学报*, 2010, 30(11): 43-46.
- [20] Yu C. Information extraction and biomass estimation of main landscape tree species in forest park based on the WorldView-2 [D]. Nanjing: Nanjing Forestry

- University, 2017.
- 余超. 基于 WorldView-2 的森林公园主要风景树种信息提取及生物量估测 [D]. 南京: 南京林业大学, 2017.
- [21] Yan J W, Chen H D, Liu L. Overview of hyperspectral image classification [J]. *Optics and Precision Engineering*, 2019, 27(3): 680-693.
- 闫敬文, 陈宏达, 刘蕾. 高光谱图像分类的研究进展 [J]. *光学精密工程*, 2019, 27(3): 680-693.
- [22] Ma J C, Zhao H D, Yang D X, et al. Design and optimization of deep convolutional neural network for aircraft target classification [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231006.
- 马俊成, 赵红东, 杨东旭, 等. 飞机目标分类的深度卷积神经网络设计优化 [J]. *激光与光电子学进展*, 2019, 56(23): 231006.
- [23] Zhang X N, Zhong X, Zhu R F, et al. Scene classification of remote sensing images based on integrated convolutional neural networks [J]. *Acta Optica Sinica*, 2018, 38(11): 1128001.
- 张晓男, 钟兴, 朱瑞飞, 等. 基于集成卷积神经网络的遥感影像场景分类 [J]. *光学学报*, 2018, 38(11): 1128001.
- [24] Liu D W, Han L, Han X Y. 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.
- [25] Li Y, Zhang H K, Xue X Z, et al. Deep learning for remote sensing image classification: a survey [J]. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018, 8(6): e1264.
- [26] Wegner J D, Branson S, Hall D, et al. Cataloging public objects using aerial and street-level images: urban trees [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 6014-6023.
- [27] Zou X H, Cheng M, Wang C, et al. Tree classification in complex forest point clouds based on deep learning [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(12): 2360-2364.
- [28] Onishi M, Ise T. Automatic classification of trees using a UAV onboard camera and deep learning [J]. *Onboard Camera*. 2018, 12(6): 15-19.
- [29] Teng W X, Wen X R, Wang N, et al. Tree species classification and mapping based on deep transfer learning with unmanned aerial vehicle high resolution images [J]. *Laser & Optoelectronics Progress*, 2019, 56(7): 072801.
- 滕文秀, 温小荣, 王妮, 等. 基于深度迁移学习的无人机高分影像树种分类与制图 [J]. *激光与光电子学进展*, 2019, 56(7): 072801.
- [30] Jing L H, Cheng Q M. Two improvement schemes of PAN modulation fusion methods for spectral distortion minimization [J]. *International Journal of Remote Sensing*, 2009, 30(8): 2119-2131.
- [31] Jing L, Hu B, Li J, et al. Automated tree crown delineation from imagery based on morphological techniques [J]. *IOP Conference Series: Earth and Environmental Science*, 2014, 17: 012066.
- [32] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [33] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [34] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1-9.
- [35] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [36] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [37] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.