

基于深度迁移学习的无人机高分影像树种分类与制图

滕文秀^{1,2}, 温小荣^{1,2*}, 王妮^{3,4}, 施慧慧³

¹南京林业大学南方现代林业协同创新中心, 江苏 南京, 210037;

²南京林业大学林学院, 江苏 南京, 210037;

³滁州学院地理信息与旅游学院, 安徽 滁州 239000;

⁴安徽省地理信息智能感知与服务工程实验室, 安徽 滁州 239000

摘要 提出一种基于深度迁移学习的无人机高分影像树种分类与制图方法。利用 ImageNet 上训练的大型卷积神经网络提取树种影像特征, 采用全局平均池化压缩树种影像特征, 使用简单线性迭代聚类生成超像素, 以超像素为最小分类单元, 生成树种专题地图。实验结果表明, 在类间差距小、类内差距大的情况下, 与小型卷积神经网络相比, 本文方法收敛更快, 总体精度和 Kappa 系数分别提高了 9.04% 和 0.1547, 超像素树种制图边界更加精确。

关键词 遥感; 树种分类; 深度迁移学习; 卷积神经网络; 超像素分割

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.072801

Tree Species Classification and Mapping Based on Deep Transfer Learning with Unmanned Aerial Vehicle High Resolution Images

Teng Wenxiu^{1,2}, Wen Xiaorong^{1,2*}, Wang Ni^{3,4}, Shi Huihui³

¹Co-Innovation Center for the Sustainable Forestry in Southern China,

Nanjing Forestry University, Nanjing, Jiangsu 210037, China;

²College of Forest, Nanjing Forestry University, Nanjing, Jiangsu 210037, China;

³School of Geographic Information and Tourism, Chuzhou University, Chuzhou, Anhui 239000, China;

⁴Anhui Engineering Laboratory of Geographical Information Intelligent Sensor and Service,

Chuzhou, Anhui 239000, China

Abstract A tree species classification and mapping method is proposed based on the deep transfer learning with unmanned aerial vehicle high resolution images. The image features of tree species are extracted using a large convolution neural network trained on ImageNet. The features of tree species images are compressed by the global average pooling. A simple linear iterative clustering method is used to generate the super-pixel, which are used as the minimum classification unit to generate tree species maps. The experimental results show that the proposed method can accelerate the convergence of the training process. The overall accuracy and Kappa coefficient are increased by 9.04% and 0.1547, respectively, compared with the small convolutional neural network method in the case of small inter-class gap and the large intra-class gap, and the boundary of the super-pixel tree mapping is more accurate.

Key words remote sensing; tree species classification; deep transfer learning; convolution neural network; super-pixel segmentation

OCIS codes 280.3640; 280.4750; 100.6890

收稿日期: 2018-10-08; 修回日期: 2018-10-17; 录用日期: 2018-10-22

基金项目: 国家重点研发计划(2016YFC0502704)、国家自然科学基金(41601455)、安徽高校省级自然科学研究重点项目(KJ2016A531)

* E-mail: wenxiu_teng@163.com

1 引言

树种分类在森林资源管理和监测^[1-2]、森林扰动评估^[3]和碳储存估算^[4]等方面具有重要价值。传统树种调查方法主要依赖于实地调查或利用大比例尺航片人工判读^[5-6],这些方法浪费人力、物力、财力,且效率低下。以卫星为代表的遥感影像树种分类方法,其效果和效率比地面调查的植被识别方法虽有明显改善,但卫星或航空传感器极易受云雾天气等外界因素的干扰,且卫星遥感运行高度较高,仅适合大范围的树种粗略分类,在小范围、精度要求较高的树种分类方面有着明显的局限性。无人机(UAV)具有灵活、高效、低成本、高空间分辨率和时间分辨率的特点^[7],可弥补卫星遥感运行周期长、费用高,以及受天气等外界因素影响大的缺点,为树种分类提供了新途径。

现有的树种分类方法主要将手工提取特征与传统机器学习算法相结合。Le Louarn 等^[8]利用光谱和纹理特征结合随机森林(RF)和支持向量机(SVM)进行树种分类;Dian 等^[9]利用冠层高度模型(CHM)、最小噪声变换及增强植被指数(EVI)结合支持向量机进行城市树种分类;任冲等^[10]综合利用光谱、纹理、植被指数、时相、地形等特征,在不同层次上采用适用于该层的支持向量机、人工神经网络(ANN)、多分类器组合(MCC)等分类方法,最后合并各层分类结果制作林地类型专题图。这些方法的表现结果很大程度上取决于人为选择的特征是否合理,而特征选择往往依靠经验,具有一定的盲目性。

深度学习^[11-12]是人工智能的重要进展之一,颠覆了传统算法的设计思路,能够自动从数据中学习特征,避免了传统算法中人工设计、提取特征的复杂性和局限性,在没有人为干预的情况下,具有极大的特征学习潜力。在遥感领域中,深度学习,特别是卷积神经网络(CNN)^[13-14],已经被广泛应用于遥感图像的分类、提取、识别和检索,全面超越了传统方法,使自动制作专题图成为可能^[15]。但关于深度学习遥感分类的研究普遍存在训练集过小的情况,由于训练集过小,这些研究大多是针对特定数据、特定任务设计的小型卷积神经网络,其网络泛化能力较弱,应用到其他场景的错误率可能会较高。从本质上解决该问题的方法有两种:一是建立针对遥感图像的大型标签数据库,包括足够多的类别以及样本,但是由于遥感影像的特殊性,这比收集千万级别的自然图像 ImageNet^[16]更加困难;二是利用迁移学习^[17],充分利

用现有的知识,将之前学习到的知识迁移到遥感影像分类任务中。在利用深度学习进行遥感影像制图方面,现有的主流技术主要分为两种:一是以固定窗口遍历整个遥感影像,识别窗口内的地物类别,并将结果作为中心像元的类别,从而实现像素级别的分类。无人机遥感影像具有很高的空间分辨率,但使用像素级分类容易出现“椒盐现象”^[18],而且这种方法效率非常低。二是通过均匀分解法将大的遥感影像平均分割成小的固定尺寸的处理单元,然后识别窗口内的地物类别,并将分类结果作为窗口内所有像素的类别。该方法与第一种方法相比虽然效率提高了,但是会产生块状地图,导致地图边界模糊^[19]。

综上所述,本文充分考虑现有树种分类方法以及深度学习遥感分类方法中存在的问题,提出一种基于深度迁移学习的无人机高分影像树种分类与制图方法,充分利用现有的数据和模型资源,将在现有数据集上学习到的知识迁移到无人机高分影像树种分类中,以减少训练时间,提高收敛速度,控制过拟合,提高模型的泛化能力。引入超像素分割技术,充分考虑像素之间的空间和光谱相关性,以超像素为基本分类单元,提高制图的精度与效率。本文为无人机高分影像树种分类与制图提供一种新的思路。

2 基础理论与方法

2.1 基础理论

2.1.1 卷积神经网络

卷积神经网络是一种采用卷积运算代替一般矩阵乘法的特殊前馈神经网络,经典的 CNN 结构由输入层、卷积层、降采样层(池化层)、全连接层和输出层组成,具体结构如图 1 所示,它具有稀疏连接、参数共享和池化三个明显特征。对人类神经系统的研究发现,人们对事物的认知过程是先通过对部分区域的认知得到局部信息,然后再对局部信息进行汇总得到全局信息。CNN 模拟人的神经网络,利用稀疏连接构建图像的局部感知野,利用网络层之间的局部空间相关性将相邻层的神经元与其相近的上层神经元连接,减少参数规模。CNN 通过参数共享,将卷积层中每一个卷积核重复地作用于整个感受野中,对图像进行卷积运算,提取图像局部特征,来减少网络各层之间的参数个数,降低过拟合风险。同时,CNN 采用降采样操作对卷积层输出的特征图进行聚合统计,实现从高层次到低层次的特征表达,简化模型复杂度,减少模型参数,降低特征映射对平移及其他形式变形的敏感度。

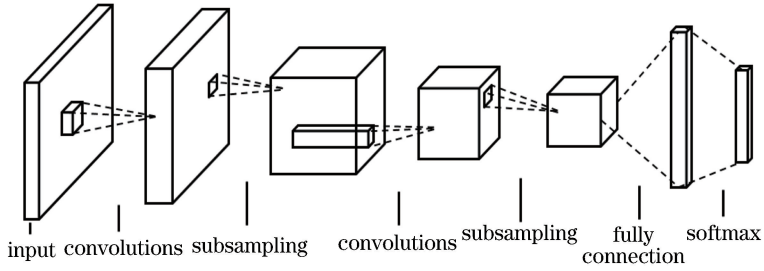


图1 卷积神经网络的模型结构

Fig. 1 Model structure of convolutional neural network

2.1.2 迁移学习

迁移学习是一种利用来自相似域中的数据解决目标域中数据不足的方法,其定义为,给定一个有标记的源域 $D_s = \{x_i, y_i\}_{i=1}^n$ 和一个无标记的目标域 $D_t = \{x_j\}_{j=n+1}^{n+m}$, 式中 x 表示影像, y 表示影像标签, i 表示具体源域影像, j 表示具体目标域影像。迁移学习的目的就是要借助 D_s 的知识,来学习目标域 D_t 的知识^[20]。传统机器学习与迁移学习的区别如图2所示:传统机器学习方法从头训练每一个任务,而迁移学习则希望通过在过去任务中学习到的知识

帮助目标任务。传统机器学习方法无论不同任务之间是否存在相似,均针对每个任务从零开始构建一个特定模型,这样孤立地解决问题,没有充分利用过去的知识,浪费了大量时间精力,对过去学习到的知识也是一种浪费;而迁移学习则是考虑任务之间有一定的相关性,通过对过去任务中学习到的知识经过微小变换甚至不经过任何改动应用到新任务中,有利于在目标任务数学模型构建中充分利用过去学到的知识,减少重复劳动和对目标任务训练数据的依赖。

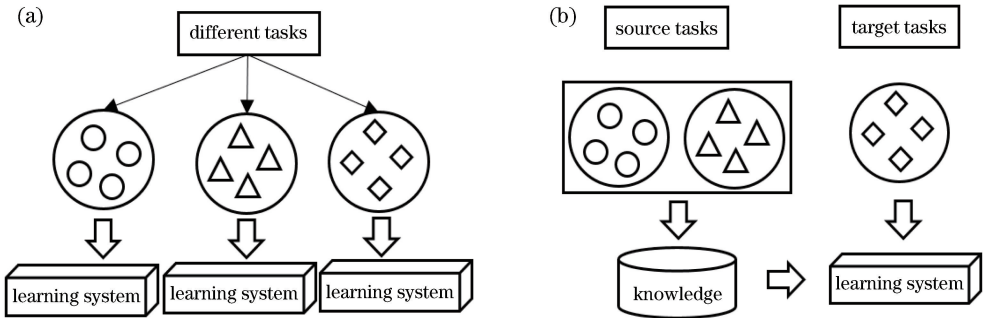


图2 传统机器学习和迁移学习对比。(a)传统机器学习;(b)迁移学习

Fig. 2 Comparison between traditional machine learning and transfer learning. (a) Traditional machine learning;

(b) transfer learning

2.2 研究方法

针对现有树种分类方法采用人工提取特征具有盲目性和不可迁移性的问题,提出一种基于深度迁移学习的无人机高分影像树种分类与制图方法。该方法首先将 ImageNet 上训练的网络参数作为大型卷积神经网络模型的初始参数,并去掉模型的全连接层和分类层,将卷积神经网络模型作为一个特征提取器提取影像的特征;然后利用全局平均池化压缩特征并与全连接层每个节点连接,采用反向传播和 Adam 优化算法,通过少量带标签的树种数据训练一个树种分类模型;最后采用超像素分割技术分割神经网络模型在其他任务上同样能表现得

更好。割整个无人机高分影像,以超像素为基本分类单元,通过训练好的树种分类模型分类所有超像素,生成树种专题地图。流程图如图3所示,其中 UAV 表示无人机。

2.2.1 深度迁移学习树种分类模型

现有的基于深度学习的遥感影像分类受限于训练数据的不足,主要针对特定问题、特定任务设计一个小型的卷积神经网络模型,这样的模型泛化能力较弱,将其应用到不同任务的遥感影像分类中可能导致错误率很高。最近的研究表明,在目前世界上最大的图像识别数据库 ImageNet 上性能更好的卷积^[21],而且卷积神经网络具有可迁移性^[22]。深度

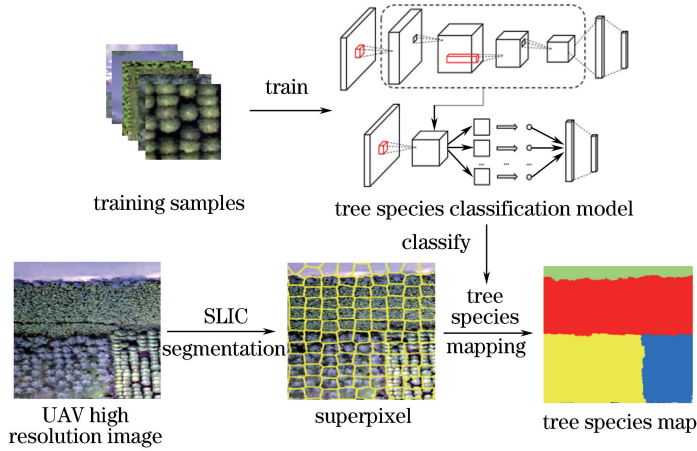


图 3 基于深度迁移学习的 UAV 高分影像树种分类与制图流程

Fig. 3 Flowchart of tree species classification and mapping of UAV high resolution image based on deep migration learning

迁移网络要比随机初始化权重效果好,并且可以加速网络的学习和优化,因此,本文将深度学习和迁移学习结合起来,将在 ImageNet 上性能较好的大型卷积神经网络模型作为本文的预训练模型,使用经训练的参数初始化网络权重,充分利用现有的数据和模型资源,将现有的在数据集上学习到的知识迁移到无人机高分影像树种分类中。本文设计的深度迁移学习模型整体框架如图4所示,该分类模型框

架主要包括两个部分:第一部分是预训练模型去掉全连接层和分类层,并使用在 ImageNet 上训练的网络参数初始化网络权重,将卷积神经网络的卷积层和池化层作为一个特征提取器提取影像特征;第二部分包括一个全局平均池化层、一个全连接层,以及一个 Softmax 分类器,以预训练模型提取的树种影像特征和树种类别标签作为输入,具体描述如下。

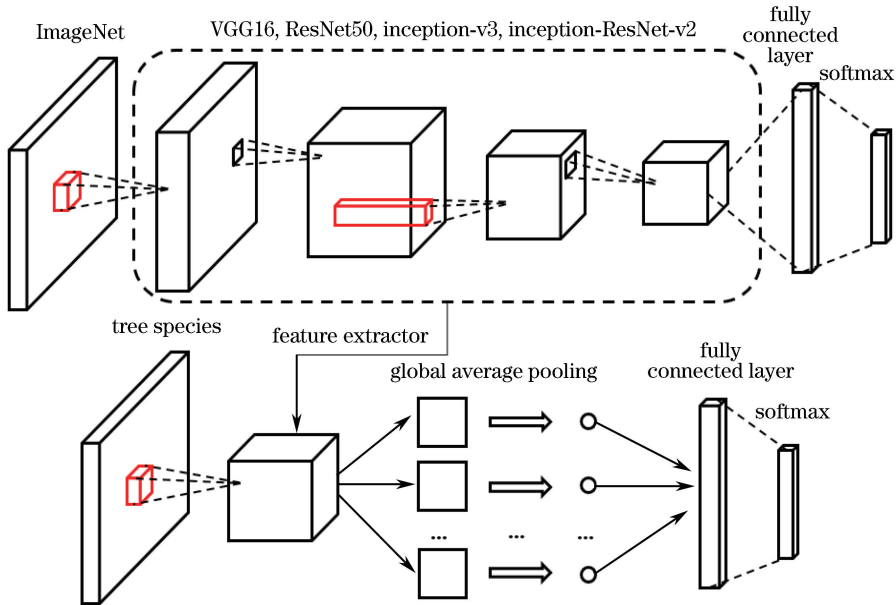


图 4 树种分类模型

Fig. 4 Tree species classification model

1) 特征提取器。将大型卷积神经网络模型去掉全连接层和分类层,使用 ImageNet 上训练的网络参数初始化网络权重,利用预训练模型的卷积和池化层提取树种影像的特征图。假设输入树种影像

为 I , F_c 表示卷积神经网络第 c 层的特征图 ($F_0 = I$),卷积和池化操作分别可表示为

$$F_c = f(F_{c-1} \otimes W_c + b_c), \quad (1)$$

$$F_c = \text{subsampling}(F_{c-1}), \quad (2)$$

式中: W_c 表示第 c 层卷积核的权值向量; \otimes 代表卷积核与第 $c-1$ 层图像或者特征图进行卷积操作; 卷积的输出与第 c 层的偏移向量 b_c 相加, 最终通过非线性的激活函数 $f(x)$ 得到第 c 层的特征图 F_c 。(2)式中 subsampling 为下采样(池化)函数。

2) 全局平均池化层。在全连接层之前添加一个全局平均池化层, 将特征提取器提取树种影像的每个特征图的所有值融合为一个特征值, 使特征维数等于特征提取器最后一层卷积层的特征图的个数, 减少全连接层参数。例如, 在 VGG16 模型中, 第一个全连接层 FC1 有 4096 个节点, 上一层卷积层输出的是 512 个 7×7 的特征图, 一共包含 $512 \times 7 \times 7 = 25088$ 个节点, 将该层与全连接层连接具有 4096×25088 个权值, 参数数量很多, 容易出现过拟合问题, 而且模型训练速度较慢; 本文方法将每个特征图融合为一个特征值, 将最后的卷积层输出的特征图融合为 512 个特征值, 然后再与全连接层连接, 则有 4096×512 个权值, 大大减少了全连接层的参数数量, 提高了训练效率, 缓解了过拟合问题。假设特征提取器最后一层卷积层的特征图大小为 $z \times v$, 式中第 k 张特征图的值用 $x_{pq}^{(k)}$ 表示, 其中 p 和 q 表示像素的位置, 经过全局池化操作后, 该特征图对应的特征值 $F^{(k)}$ 可表示为

$$F^{(k)} = \frac{1}{m \cdot n} \left[\sum_{p=0}^z \sum_{q=0}^v x_{pq}^{(k)} \right] \quad (3)$$

3) 全连接层。经过特征提取和全局平均池化后, 将全局平均池化后的所有特征值与全连接层每个节点连接。本文设计的全连接层包含 1024 个神经元, 使用修正线性单元 (Relu) 作为激活函数, 可表示为

$$\text{Relu}(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases} \quad (4)$$

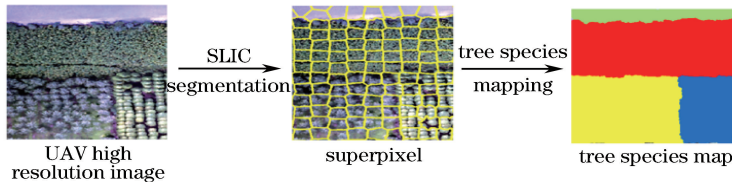


图 5 树种制图流程图

Fig. 5 Flowchart of tree species mapping

首先对遥感影像进行超像素分割。超像素分割采用简单线性迭代聚类 (SLIC) 算法^[23]生成。超像素分割同时考虑像素的空间和光谱属性, 具有良好的内部一致性, 并且每个超像素具有相似的大小, 适合作为卷积神经网络的输入。具体算法在 $[l, a, b,$

4) Softmax 分类器。最后一层使用 Softmax 分类器, 将网络输出的特征放到 Softmax 分类器中进行分类, 其类标签为 $y \in \{1, 2, \dots, K\}$, K 为类别标签数量, 本文 $K = 4$, 分为杨树、水杉、竹子、银杏 4 类。

对于模型的训练, 传统方法采用随机梯度下降 (SGD) 进行反向传播 (BP), 逐层更新网络的可训练参数, 但该方法容易收敛到局部最优, 并且学习率选择比较困难。Adam 优化算法将动量与 RMSProp 优化算法相结合, 它利用梯度的一阶矩估计和二阶矩估计自适应选择每个训练参数的学习率, 经过偏置校正后, 每一次迭代学习率都有个确定范围, 使得参数比较平稳, 并且该算法对内存需求较小, 因此, 本文采用 Adam 优化算法和反向传播更新模型参数, 具体表示为

$$w_t = w_t - 1 - a_t m_{dw} / (\sqrt{\hat{v}_{dw}} + \epsilon) \quad (5)$$

$$b_t = b_t - 1 - a_t m_{db} / (\sqrt{\hat{v}_{db}} + \epsilon) \quad (6)$$

式中: w_t, b_t, a_t 分别为第 t 次迭代时的权值、偏置、学习速率; \hat{m}_{dw} 和 \hat{m}_{db} 分别为权值与偏置的偏差纠正动量值; $\sqrt{\hat{v}_{dw}}$ 和 $\sqrt{\hat{v}_{db}}$ 分别为权值与偏置的偏差纠正 RMSProp 值; ϵ 为 Adma 算法中的超参数。

2.2.2 超像素分割树种制图

现有的基于卷积神经网络的遥感影像制图方法中, 滑动窗口方法没有考虑像素的空间和光谱相关性, 地图边界模糊; 而像素级别的分类效率较低, 容易出现“椒盐现象”, 已经被证明不适于高空间分辨率遥感影像制图。为此, 本文引入超像素分割技术, 首先对整幅影像进行超像素分割, 以超像素为最小分类单元作为模型的输入, 实现整幅遥感影像的分类, 最终生成树种专题地图。制图流程如图 5 所示。

$x, y]$ 五维空间实现, 式中 $[l, a, b]$ 是像素点在 lab 颜色空间对应的像素值, l 表示亮度, a 表示从洋红色至绿色的范围, b 表示从黄色至蓝色的范围, $[x, y]$ 是像素点在图像平面的位置坐标。该算法的核心步骤如下。

1) 确定聚类中心个数 u 。 u 个聚类中心最终生成 u 个超像素,采用过分割的方式对无人机高分影像进行超像素分割, u 值的具体确定方式为

$$u = \frac{m}{s_1/r} \times \frac{n}{s_2/r}, \quad (7)$$

式中: m 和 n 为无人机高分影像大小; r 为无人机高分影像的分辨率,单位为 m ; s_1 和 s_2 为超像素样地实际大小,单位为 m 。

2) 初始化聚类中心。假设无人机高分影像像素个数为 N ,在影像上按照等距离 L 选取 u 个初始聚类中心,式中 $L = \sqrt{N/u}$ 。 L 取 $\sqrt{N/u}$ 是为保证生成的超像素块形状规则,分布均匀且紧凑,生成的超像素更适合卷积神经网络的输入。

3) 聚类中心校正。为了避免聚类中心处于边界点或者奇异点,计算原聚类中心周围 3×3 邻域内所有像素点的梯度值,将聚类中心移动到梯度值最小的位置。

4) 超像素分割。搜索聚类中心 $2s \times 2s$ 范围的邻域,将相似度大于聚类阈值的像素与该中心聚类。迭代聚类过程直至收敛,完成整幅无人机高分影像分割。对于相似度的度量,虽然欧氏距离在国际照明委员会(CIELAB)颜色空间中的小距离计算中具有直观的意义,但如果空间上点的距离超过了颜色距离的极限,则会使欧氏距离偏向于像素颜色的相似性,最终导致超像素不再考虑区域边界,为此本文利用 lab 色彩空间对应的值和 x 、 y 坐标值来衡量像素间的相似度,具体表示为

$$d_{lab} = \sqrt{(l_g - l_h)^2 + (a_g - a_h)^2 + (b_g - b_h)^2}, \quad (8)$$

$$d_{xy} = \sqrt{(x_g - x_h)^2 + (y_g - y_h)^2}, \quad (9)$$

$$H_s = d_{lab} + \frac{o}{s} d_{xy}, \quad (10)$$

式中: d_{lab} 为像素点 g 、 h 的色彩差异度; d_{xy} 为像素点 g 、 h 的欧氏距离; H_s 为聚类阈值; s 是聚类中心的距离; o 为紧凑因子,取值范围为 $[1, 40]$ 。当 o 较大时,产生的超像素更紧凑(面积与周长比较低);当 o 较小时,产生的超像素更紧密地附在图像边界上,但其大小和形状不那么规则。为了使产生的超像素形状相对规则适应卷积神经网络的输入以及树种专题图边界相对明显,本文 o 取值为 10。

超像素分割后,利用超像素分割结果的外包矩形裁剪影像,将裁剪后的影像作为卷积神经网络的输入进行分类,分类结果作为超像素边界内的所有像素的类别,完成所有超像素的分类,合并分类结果

生成树种专题地图。

3 实验与分析

为验证本文方法的有效性,以无人机高分影像数据作为数据源,采用本文方法进行树种分类与制图实验,并评价其有效性。在树种分类方面,利用混淆矩阵、总体精度(OA)和 Kappa 系数来评价本文树种分类方法的性能,探讨不同预训练模型对树种分类精度的影响,并和小型卷积神经网络方法进行对比,探讨使用 ImageNet 训练网络初始化树种分类模型网络权重的适应性。在树种制图方面,采用超像素分割和均匀分解法进行树种制图,利用兴趣点视觉解释评价制图效果。

3.1 实验数据与实验环境

以江苏省东台市东台林场作为研究区。东台林场位于江苏省东台市,地理坐标介于 $120^\circ 47' 11'' \sim 120^\circ 52' 0'' E$ 、 $32^\circ 53' 30'' \sim 32^\circ 51' 17'' N$,无人机高分影像获取于 2016 年 8 月上旬,选择天气晴朗、云量较少的天气进行航拍,采用的无人机为数字绿土八旋翼无人机,无人机搭载了该公司研发的 Liortho 高分辨率影像获取系统[由高分辨率相机、控制系统、差分全球定位系统(GPS)或 POS (Position)系统组成],相对航拍高度为 200 m。将获取的单幅影像利用 Li Mapper 软件进行处理,使用自动空中三角测量和光束平差法原理自动提取影像特征点,计算正确的位置参数,并进行几何校正、正射校正,最后自动拼接成一幅影像,其影像分辨率为 0.15 m。从东台林场无人机高分影像中选择一块具有代表性的样地作为实验区域,影像大小为 $2261 \text{ pixel} \times 2913 \text{ pixel}$,研究区域中包括杨树、水杉、竹子、银杏 4 类,共采集样本 806 个,样本大小为 $100 \text{ pixel} \times 100 \text{ pixel}$,其中,水杉、杨树、竹子、银杏样本分别为 475、170、70、91 个,随机选择其中的 20% 作为训练样本,其余 80% 作为测试样本,树种样本如图 6 所示。从图中能够明显看出,树种影像受树密度、树龄、树种、光照、背景等影响,类间差距小、类内差距大,并且树种训练样本较少,分类难度较大。

实验环境为 Windows 10 64bit 操作系统,CPU 为 Core i7-8700K @ 3.70 GHz,16 GB 内存,GPU 为 Nvidia TITAN Xp 12 GB 显存。深度学习平台采用 TensorFlow 1.8.0、Keras 2.2.0,栅格和矢量数据处理采用 GDAL/OGR 2.2.4,数据可视化工具采用 Tensorboard 1.8.0、Matplotlib 2.0.2、sklearn 0.19.0,编程语言采用 Python 3.6.3。

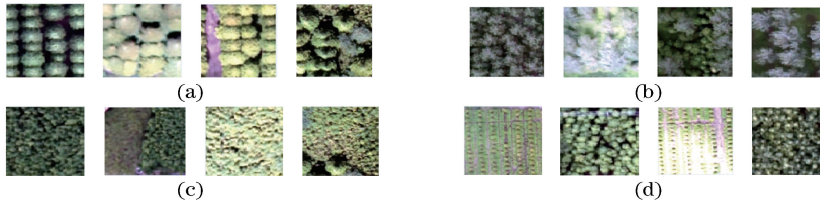


图6 树种样本。(a)水杉;(b)杨树;(c)竹子;(d)银杏

Fig. 6 Tree species samples. (a) Metasequoia; (b) populus; (c) bamboo; (d) ginkgo

3.2 不同预训练模型分类对比

2012年, Hinton 课题组提出的 AlexNet 在 ImageNet 图像分类竞赛中以准确度超越第二名 11% 的巨大优势夺得冠军^[24], 引爆人工智能领域。随后更多的卷积神经网络模型被相继提出, 更深更复杂的模型不断刷新着准确率。从网络结构看, 卷积神经网络从早期的无分支模型, 即首尾相接、没有平行层的模型, 如 AlexNet、VGG16、VGG19, 演化为模块堆叠模型, 在模型深度和宽度上都有增加, 如 GoogLeNet、ResNet50、Inception-v3, 到现在的 Inception-ResNet-v2, 在 Inception-v3 模型基础上, 结合 ResNet50 深度模型中残差连接思想构建混合模型。为探讨不同模型对树种分类的影响, 本文选择 4 个比较有代表性的网络作为预训练模型, 具体包括 VGG16、ResNet50、Inception-v3 和 Inception-ResNet-v2。训练模型前, 将带标签的树种影像缩放到模型需要的大小, 其中 VGG16 和 ResNet50 输入图像大小为 224 pixel × 224 pixel, Inception-v3 和 Inception-ResNet-v2 输入图像大小为 229 pixel × 229 pixel。对于网络训练, 为了公平起见, 4 个模型使用相同的超参数, 速率衰减周期设置为 100 epoch, 所有训练样本完整输入模型 1 次为 1 epoch, Batch size (每次输入模型的样本数量) 设置为 32。对于 Adam 优化算法, 将初始学习速率设置为 0.0001, 速率衰减因子设置为 0.9, epsilon 设置为 10^{-8} 。每个模型的训练精度变化和损失变化如图 7 所示, 从图中可以看出, 4 个模型在训练 20~30 epoch 后精度损失基本稳定, 其中 VGG16 和 ResNet50 具有更高的精度和更好的稳定性, 损失也降得更低, 其中 VGG16 无论是精度、稳定性还是损失, 在 4 个模型中都是表现最好的。

各模型的分混淆矩阵、OA 和 Kappa 系数如图 8 所示。VGG16 总体分类精度为 96.11%, Kappa 系数为 0.9338; ResNet50 总体分类精度为 93.30%, Kappa 系数为 0.8852; Inception-v3 总体分类精度为 89.88%, Kappa 系数为 0.8235; Inception-ResNet-v2 总体分类精度为 91.12%, Kappa 系数为

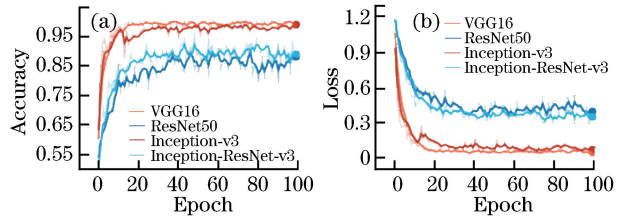


图7 不同模型的训练精度及损失曲线。(a)训练精度曲线; (b)训练损失曲线

Fig. 7 Training accuracy and loss curves of different models.

(a) Training accuracy curves; (b) training loss curves

0.8475。从图中可以看出, VGG16 和 ResNet50 分类精度明显高于 Inception-v3 和 Inception-ResNet-v2, 其中 VGG16 精度最高, 这与 ImageNet 数据集上的表现不一致。本实验中, 在 ImageNet 数据集上表现更好的 Inception-v3 和 Inception-ResNet-v2 模型并没有取得最优的结果, 而 VGG16 模型在训练和测试上表现最好。这可能是因为树种影像和自然图像不同, 不具有非常复杂的语义特征, 因此, 相对较小的卷积神经网络反而取得了更好的效果。故将 VGG16 作为树种分类模型的预训练模型。

3.3 与小型卷积神经网络分类方法对比

很多研究已证明, 卷积神经网络的性能超越传统的机器学习算法, 因此, 本文对比实验部分不再重复讨论卷积神经网络和传统方法相比的优势, 而是与一个小型的卷积神经网络进行对比, 并探讨使用 ImageNet 训练网络初始化树种分类模型网络权重的适应性。小型卷积神经网络包括 3 个卷积层、3 个池化层和 1 个全连接层, 为了公平起见, 采用与本文方法同样的超参数。图 9 给出了两种方法的训练精度和损失变化曲线。小型卷积神经网络的精度一开始相对较低, 在经过 30 epoch 后保持稳定; 而 VGG16 初始精度相对较高, 经过 20 epoch 后保持稳定, 说明本文方法收敛速度更快, 并且在相对稳定的情况下精度更高。小型卷积神经网络的损失一开始很高, 经过 20 epoch 后才保持稳定, 而 VGG16 初始损失相对较低, 并且在经过 10 epoch 后就保持稳

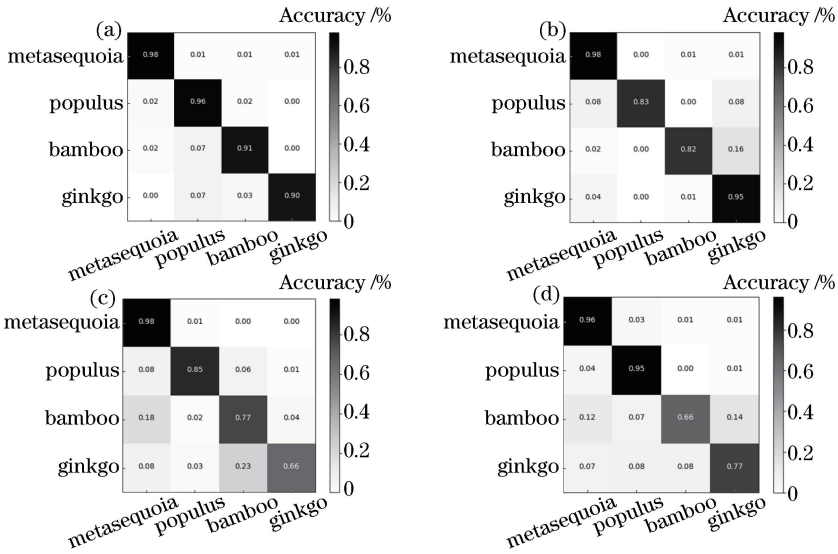


图 8 不同模型分类的混淆矩阵。(a) VGG16; (b) ResNet50; (c) Inception-v3; (d) Inception-ResNet-v2

Fig. 8 Confusion matrices of classification with different models. (a) VGG16; (b) ResNet50;

(c) Inception-v3; (d) Inception-ResNet-v2

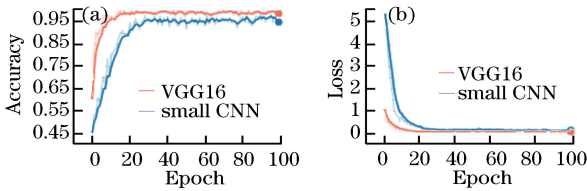


图 9 不同方法的训练精度及损失曲线。

(a) 训练精度曲线; (b) 训练损失曲线

Fig. 9 Training accuracy and loss curves of different methods.

(a) Training accuracy curves; (b) training loss curves

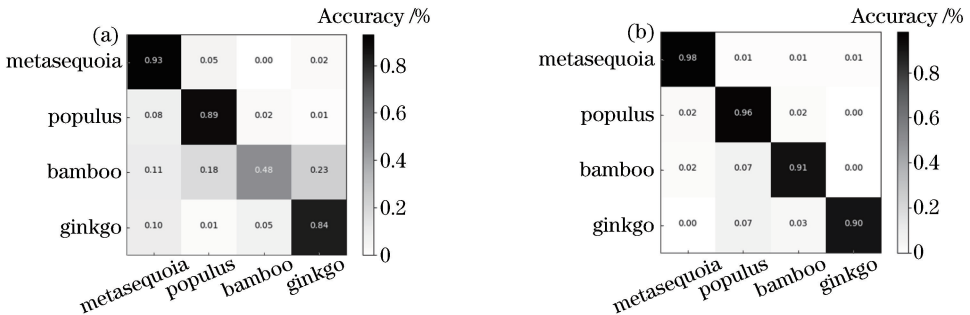


图 10 不同方法分类的混淆矩阵。(a) 小型卷积神经网络; (b) 本文方法

Fig. 10 Confusion matrices of classification with different methods. (a) Small CNN; (b) proposed method

3.4 超像素分割树种制图

树种分类模型训练好后,使用超像素分割和均匀分解法得到基本分类单元,利用树种分类模型对超像素单元进行分类,最后合并分类结果生成树种专题地图,其中,超像素样地实际大小为 12 m×12 m,均匀分解法利用 100 pixel×100 pixel窗口裁剪整个影像,制图结果如图 11 所示。为了更有效地比较两种方法,使用原图叠加

定,并且损失降得更低。图 10 给出了小型卷积神经网络和本文方法的分类混淆矩阵。小型卷积神经网络总体分类精度为 87.07%,Kappa 系数为 0.7791;本文算法总体分类精度为 96.11%,Kappa 系数为 0.9338,比小型卷积神经网络分别高 9.04%和 0.1547。综上所述,使用预训练模型初始化网络参数与重新训练一个小型的卷积神经网络相比具有更快的收敛速度和更高的分类精度,是一种更有效的树种分类方法。

树种专题地图显示制图结果,树种专题地图的透明度为 30%,并同时放大地图中 3 个兴趣点来对比两种方法的细节差异,如图 11 中 1、2、3 区域分别包含水杉和竹子、杨树和水杉、银杏和杨树。通过对比两种方法的放大区域可以看出,由于超像素分割充分考虑了像素之间的空间和光谱相关性,超像素分割和均匀分解法相比树种边界更加精确,是一种更有效的制图方法。

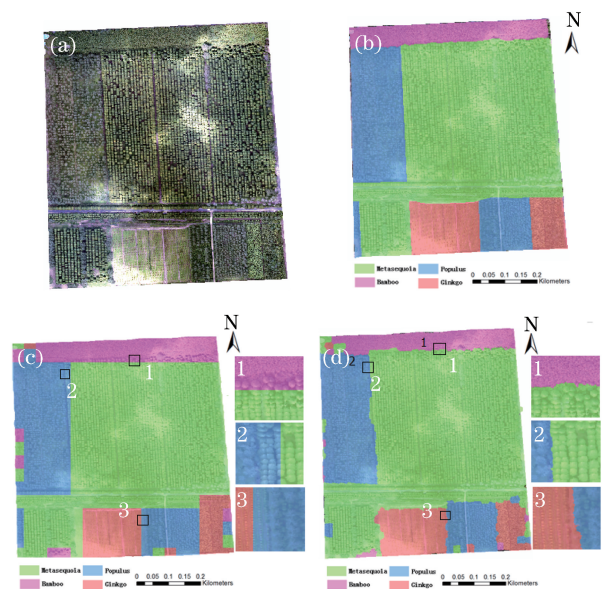


图 11 不同方法的树种地图。(a)原始影像;(b)实际地图;
(c)均匀分解法;(d)本文方法

Fig. 11 Tree species maps with different methods. (a) Original image; (b) actual map; (c) uniform decomposition method; (d) proposed method

4 结 论

通过研究树种分类及卷积神经网络在遥感影像分类应用中存在的问题,提出了基于深度迁移学习的无人机高分影像树种分类与制图方法,将深度学习和迁移学习结合起来,充分利用现有知识,将现有知识迁移到无人机高分影像树种分类中,实现了端到端的树种分类,避免了人工提取特征的不可迁移性,解决了训练大型卷积神经网络树种标签不足容易出现过拟合及小型卷积神经网络泛化能力不足的问题。实验结果表明,和小型的卷积神经网络相比,本文方法在树种影像受到树密度、树龄、树种、光照、背景等影响,类间差距小、类内差距大、树种标签较少、分类难度较大的情况下,具有收敛更快,精度更高的特点,其总体精度和 Kappa 系数分别提高 9.04% 和 0.1547,同时能满足小样本分类需求。本文方法充分考虑像素的空间和光谱相关性,相比均匀分解法,树种边界更加精确。将深度学习应用到无人机高分影像树种分类中,为树种分类研究提供了一个新的思路。

参 考 文 献

[1] Plourde L C, Ollinger S V, Smith M L, *et al.* Estimating species abundance in a northern temperate

forest using spectral mixture analysis [J]. Photogrammetric Engineering & Remote Sensing, 2007, 73(7): 829-840.

- [2] Cho M A, Mathieu R, Asner G P, *et al.* Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system[J]. Remote Sensing of Environment, 2012, 125(10): 214-226.
- [3] Dale V H, Joyce L A, McNulty S, *et al.* Climate change and forest disturbances[J]. BioScience, 2001, 51(9): 723-734.
- [4] Thomas S C, Malczewski G. Wood carbon content of tree species in Eastern China: interspecific variability and the importance of the volatile fraction[J]. Journal of Environmental Management, 2007, 85(3): 659-662.
- [5] Gong P. Conifer species recognition: an exploratory analysis of in situ hyperspectral data [J]. Remote Sensing of Environment, 1997, 62(2): 189-200.
- [6] Waser L T, Ginzler C, Kuechler M, *et al.* Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS₄₀) and RC₃₀ data [J]. Remote Sensing of Environment, 2011, 115(1): 76-85.
- [7] Fan C X, Han J, Xiong Z J, *et al.* Application and status of unmanned aerial vehicle remote sensing technology [J]. Science of Surveying and Mapping, 2009, 34(5): 214-215.
范承啸, 韩俊, 熊志军, 等. 无人机遥感技术现状与应用 [J]. 测绘科学, 2009, 34(5): 214-215.
- [8] Le Louarn M, Clergeau P, Briche E, *et al.* "Kill two birds with one stone": urban tree species classification using bi-temporal Pléiades images to study nesting preferences of an invasive bird [J]. Remote Sensing, 2017, 9(9): 916.
- [9] Dian Y Y, Pang Y, Dong Y F, *et al.* Urban tree species mapping using airborne LiDAR and hyperspectral data [J]. Journal of the Indian Society of Remote Sensing, 2016, 44(4): 595-603.
- [10] Ren C, Ju H B, Zhang H Q, *et al.* Multi-source data for forest land type precise classification [J]. Scientia Silvae Sinicae, 2016, 52(6): 54-65.
任冲, 鞠洪波, 张怀清, 等. 多源数据林地类型的精细分类方法 [J]. 林业科学, 2016, 52(6): 54-65.
- [11] Chen Y, Fan R S, Wang J X, *et al.* Segmentation of high-resolution remote sensing image combining phase consistency with watershed transformation [J].

- Laser & Optoelectronics Progress, 2017, 54(9): 092803.
陈洋, 范荣双, 王竞雪, 等. 结合相位一致和分水岭变换的高分辨率遥感影像分割方法[J]. 激光与光电子学进展, 2017, 54(9): 092803.
- [12] Le Cun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [13] Shu C X, He Y T, Sun Q K. Point cloud registration based on convolutional neural network[J]. Laser & Optoelectronics Progress, 2017, 54(3): 031001.
舒程珣, 何云涛, 孙庆科. 基于卷积神经网络的点云配准方法[J]. 激光与光电子学进展, 2017, 54(3): 031001.
- [14] Yan Q, Li H, Jing L H, *et al.* Automatic extraction algorithm of seismic landslide information based on after-calamity high-resolution remote sensing image [J]. Laser & Optoelectronics Progress, 2017, 54(11): 112801.
闫琦, 李慧, 荆林海, 等. 灾后高分辨率遥感影像的地震型滑坡信息自动提取算法研究[J]. 激光与光电子学进展, 2017, 54(11): 112801.
- [15] Gong J Y, Ji S P. Photogrammetry and deep learning [J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(6): 693-704.
龚健雅, 季顺平. 摄影测量与深度学习[J]. 测绘学报, 2018, 47(6): 693-704.
- [16] Deng J, Dong W, Socher R, *et al.* ImageNet: a large-scale hierarchical image database [C] // IEEE Computer Vision and Pattern Recognition, 2009: 248-255.
- [17] Pan S J, Yang Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [18] Li L, Shu N. Object-oriented classification of high-resolution remote sensing image using structural feature [C] // International Congress on Image and Signal Processing, 2010: 2212-2215.
- [19] Zhang F, Du B, Zhang L P. Scene classification via a gradient boosting random convolutional network framework [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(3): 1793-1802.
- [20] Kermany D S, Goldbaum M, Cai W J, *et al.* Identifying medical diagnoses and treatable diseases by image-based deep learning [J]. Cell, 2018, 172(5): 1122-1131.
- [21] Yosinski J, Clune J, Bengio Y, *et al.* How transferable are features in deep neural networks? [C] // International Conference on Neural Information Processing Systems, 2014: 3320-3328.
- [22] Kornblith S, Shlens J, Le Q V. Do better imagenet models transfer better? [J]. arXiv: 1805.08974, 2018.
- [23] 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.
- [24] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems, 2012: 1097-1105.