

基于 UNet-ResNet14*半监督学习的无人机影像森林树种分类

陈龙伟¹, 周小成^{1*}, 李传昕¹, 林华章², 王永荣³, 崔永红⁴

(1. 福州大学空间数据挖掘与信息共享教育部重点实验室, 福州 350108; 2. 福建省大田桃源国有林场, 三明 366199; 3. 福建省漳平五一国有林场, 龙岩 364400; 4. 福建省林业调查规划院, 福州 350003)

摘要: 无人机遥感在森林树种精细和高效分类制图中具有巨大的潜力。为了快速准确获取森林的优势树种分布信息, 该研究探讨了半监督学习方法在树种分类方面的有效性。以福建省福州市、龙岩市和三明市的4个试验区为例, 构建精简的 ResNet18 为主干的 UNet 树种分类模型 (UNet-ResNet14*), 使用交叉熵和 Dice 系数的联合损失函数来优化模型参数, 对比分析 Self-training 和 Mean teacher 两种不同的半监督学习方法在无人机影像森林树种分类模型的泛化能力。结果表明, 以 ResNet14*作为主干的分类模型与其他模型相比精度更高且预测速度更快, 当联合损失函数权重值为 0.5 的情况下模型预测效果最好, 总体精度达到了 91.15%。经过 Self-training 的模型在木荷、马尾松、杉木 3 个样本充足的类别中精度均有所提升, 总精度为 91.08%, 比原始模型略低, 但在独立验证区的精度为 88.50%, 比原始模型高; Mean teacher 方法的总精度为 88.56%, 在独立验证区的精度为 73.56%。因此, 研究认为可以采用 Self-training 半监督方法结合 UNet-ResNet14*的方案快速得到试验区的树种组成信息。

关键词: 无人机; 遥感; 森林; 树种分类; 可见光; Unet; ResNet; 半监督学习

doi: 10.11975/j.issn.1002-6819.202310172

中图分类号: TP79;S757.2

文献标志码: A

文章编号: 1002-6819(2024)-01-0225-10

陈龙伟, 周小成, 李传昕, 等. 基于 UNet-ResNet14*半监督学习的无人机影像森林树种分类[J]. 农业工程学报, 2024, 40(1): 225-234. doi: 10.11975/j.issn.1002-6819.202310172 <http://www.tcsae.org>
CHEN Longwei, ZHOU Xiaocheng, LI Chuanxin, et al. Classification of tree species based on UNet-ResNet14* semi-supervised learning using UAV images[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(1): 225-234. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202310172 <http://www.tcsae.org>

0 引言

森林生态系统是地球碳循环重要的汇和库, 在维护全球生态平衡和保护生物多样性中发挥了重要作用。及时获取森林树种分布与林分树种比例, 能够直接帮助林业工作人员快速了解森林资源和制定合理的森林管理体系, 这对于林业资源的可持续发展尤为重要。相较于传统林业资源调查方法, 遥感技术的发展使得森林资源调查在效率上和规模上都取得了重大的突破。早期的遥感数据难以获取且成本较高, 而现在, 从航天高分辨率遥感卫星数据到近地面无人机数据都已广泛应用在树种分类领域。相较于传统的卫星监测, 无人机几乎能在全天候全天时的条件下作业, 可以更高效和更机动地获取精细的树种组成信息。随着林业业务需求的不断增加, 许多类型的传感器都能搭载在无人机上应用于森林树种分类制图。与可见光传感器相比, 多光谱^[1]、高光谱^[2-3]和机载雷达 (light detection and ranging, LiDAR)^[4-6]能够区分更多的树种类型和得到更高的树种分类精度。然而

这些数据多受制于研究区的立地条件难以适用于多变的环境, 同时数据获取和处理成本较高, 难以大面积推广使用。

可见光传感器具有高分辨率、高信噪比、成本低廉和影像数据易获取等优势^[7-9], 在森林树种分类制图中具有重要潜力^[10-11]。WANG 等^[12]在城市公园场景中基于无人机超高分辨率可见光图像提取绿化区域的树种信息, 该研究使用基于对象的随机森林分类器实现了单木级的树种分类, 总体精度达到了 91.3%, 提供了城市公园绿化场景下的树种识别方案; NATESAN 等^[13]探讨了无人机可见光影像对常见的经济针叶树种进行单木级分类的能力, 评价了 DenseNet 模型分别在 2015、2016 和 2018 年对同一研究区的分类精度, 各年份的总体精度分别为 83%、84.5% 和 83.6%, 证明了该模型在不同季节、不同的光照条件下依然具有一定的稳健性; NEZAMI 等^[14]使用 3D-CNN 模型对芬兰北方森林的 3 种树种进行分类, 该研究基于可见光航片衍生的摄影测量点云和原有的数字表面模型构建了森林树种的冠层高度模型 (canopy height model, CHM), 并分别对比了仅使用无人机红-绿-蓝 (red-green-blue, RGB) 影像和 RGB+CHM 两个不同策略下的模型精度, 最终 2 种策略总体精度分别为 97.1% 和 97.5%。

当前, 基于深度学习的技术需要使用大量的数据集来进行训练, 而这些数据集通常需要手动标注。然而,

收稿日期: 2023-10-24 修订日期: 2024-01-12

基金项目: 福建省科技厅高校产学研合作项目 (2022N5008); 福建省科技厅对外合作项目 (202210007)

作者简介: 陈龙伟, 研究方向为环境与资源遥感。Email: 825754917@qq.com

*通信作者: 周小成, 博士, 副研究员, 研究方向为环境与资源遥感。Email: zhouxc@fzu.edu.cn

在某些专业领域中,收集和制作数据集的成本较高。另外,从外部采集回来的数据通常需要专业人员进行标注,而标注人员的知识和经验水平在很大程度上决定了标签的质量。因此,通过人工标注的方式获取大规模的林业病害标注数据集通常是不现实的。在这种情况下,如何充分利用大量未经标注的数据和少量已经标注的数据来进行基于深度学习方法的森林树种分类,成为一个需要研究和解决的重要问题。

针对上述问题,本文提出了一种基于半监督深度学习的森林树种分类框架来有效利用少量带标签样本和大量未标记样本,以实现低成本且高精度的森林树种分类。针对福建省复杂的山地森林环境,本文以如何快速有效且低成本地获取试验区的树种组成为出发点,采集分布于福建省福州市、三明市、龙岩市 4 个试验区的无人机可见光影像,构建以 ResNet14*为主干的轻量化 UNet 树种分类模型,探究 Self-training 和 Mean teacher 两种不同的半监督学习方法在无人机影像中的模型泛化能力,从而实现林区的快速树种制图,以期林业部门的树种分布信息的获取与森林结构调整提供技术参考。

1 研究区与数据源

1.1 研究区概况

福建省地处中国东南沿海,属于典型的亚热带季风气候,常年阳光充足,年平均气温在 17~21℃ 之间,是中国雨量最丰富的省份之一。各个试验区均包括三类或三类以上的森林树种,同时树种组成比例均不相同,每个试验区均为针阔混交的人工林。为了验证树种分类模型的有效性,本文从福建省福州市、三明市和龙岩市选择 4 个试验区开展示范研究。

在 4 个试验区中,福州市闽侯样区 A,面积约 8.96 hm²,海拔在 130~240 m,树种以木荷、马尾松为主,其中还存在少量的杉木、竹类;福州市闽侯样区 B,面积约 10.31 hm²,海拔在 250~440 m,树种以杉木、马尾松为主,并存在少量的福建柏、桉树及其他阔叶树;样区 C 为龙岩市漳平样区,面积约 9.57 hm²,海拔在 180~420 m,树种以木荷、马尾松和杉木为主,其他阔叶树在试验区内零星分布;样区 D 为三明市大田样区,面积约 10.92 hm²,海拔在 580~700 m,树种以马尾松和杉木为主。4 个试验区都是针阔混交的人工林,均在山地丘陵地带,树龄均为 20~30 年。4 个试验区内的树种组成比例各不相同,均有少量的裸地和草地出现。试验区具体位置及数字正射影像(digital orthophoto map, DOM)如图 1 所示。

1.2 无人机数据

本文使用大疆精灵 4 RTK 版本作为数据采集平台,数据采集时间为 2021 年 5 月 26 日和 2022 年 10 月 2 日,在晴朗无云少风的情况下获取 4 个样地的无人机数据。该无人机平台搭载了 1 英寸 CMOS 图像传感器,该可见光相机具有 2 000 万像素,包括红、绿、蓝 3 个波段。考虑到研究区地形复杂,本研究采用仿地飞行的方法获

取试验区的影像数据,从而保证采集的影像中树冠足够清晰,能够区分林冠边界。4 个试验区的航线规划相关参数设置如表 1 所示。

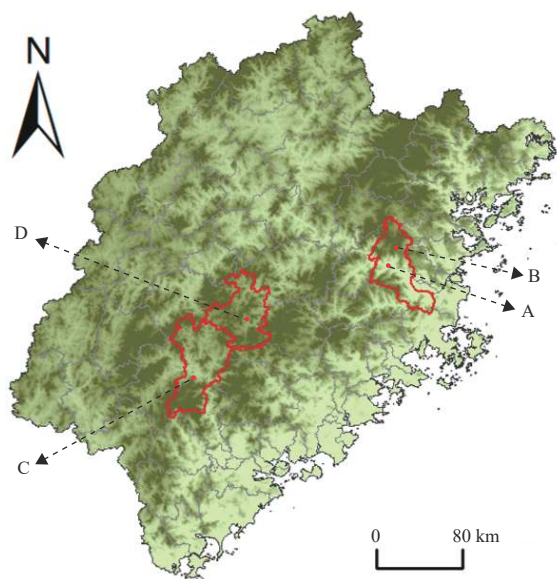


图 1 不同试验区在福建省的位置

Fig.1 Locations of different experimental areas

表 1 不同试验区的航线规划参数

Table 1 Route planning parameters for different experimental areas

参数 Parameters	A	B	C	D
拍摄时间 Shooting time	2021-05-26		2022-10-02	
分辨率 Resolution/cm	3.2	3.8	3.6	3.1
航向重叠率 Heading overlap rate/%			70	
旁向重叠率 Lateral overlap rate/%			80	
仿地飞行精度 Terrain-following flight accuracy/m			10	

在数据处理阶段,本文使用 Photoshop 软件对色调不一致的无人机航片匀光匀色,再使用 Pix4Dmapper 软件将预处理后的无人机可见光航片进行拼接生成数字正射影像和数字表面模型(digital surface model, DSM)。

1.3 地面调查数据与解译标志

在获取无人机数据的同时,同步开展了示范区树种的地面调查。首先以历史森林小班数据作为参考数据,了解研究区内的树种组成比例及大致分布;由于木荷、马尾松、杉木、竹类、福建柏 5 个树种涵盖了研究区内 90% 以上的树木,因此这 5 个树种作为本文研究的主要树种类型。研究区内的其他树种定义为其他树种;森林以外的区域还包括裸地、耕地、草地、建筑用地等地物类别,本文将森林冠层区域以外的其他地物统一归为非森林或非林冠区。此外,本文还将不同物候期的相同树种归为一个类别,以此保持最终模型在某一树种不同物候期的稳健性。最终确定的分类体系为木荷、马尾松、杉木、竹类、福建柏、其他阔叶树、非林冠 7 个类别,各类别的解译标志如图 2 所示。



图 2 不同树种在无人机可见光影像上的影像特征

Fig.2 Different categories of interpreted signs on visible images

1.4 样本集构建和数据增广

深度学习样本的质量与数量是影响模型分类精度的重要因素之一，直接决定了模型的分类准确率和分类结果的可靠性。为此本文基于 3 个试验区 A、B、C 的正射影像构建优势树种的样本集。具体流程如下：首先，采用重采样的方法将 3 个试验区的影像分辨率统一为 3 cm；其次，在重采样后的正射影像上使用 ArcGIS 软件绘制该区域的类别标签，并编写 Python 脚本将影像和类别标签进行裁剪，最终得到 3 195 张 256×256 的样本块。样本集中木荷、杉木、马尾松、竹类、福建柏、其他树种的样本数量比例约为 106:84:92:16:2:1。为了测试模型的精度表现，样本集被划分为训练集、测试集两部分，其中 80% 的样本被用于模型训练，20% 的样本被用于模型测试，最终获得标记的训练集 2 556 张、测试集 639 张；本研究将试验区 D 作为独立验证区域，其正射影像作为未标记样本集的数据源，为了将分辨率与原有数据集保持一致，同样使用双线性内插法将正射影像的分辨率重采样为 3 cm，并将重采样后的正射影像裁剪为 256×256 的影像块，去除非森林区域占比大于 50% 的样本块

后，最终得到的未标记样本集为 1 563 张，作为 Self-training 方法和 Mean teacher 方法共用的未标记的样本集。

对于树种分类任务，需要花费大量时间制作样本集，而数据增广是一种能有效提高模型泛化性能的方法。综合研究区的环境条件和数据特点，本文选择镜像、旋转、裁剪、等比例放大（缩小）、透视变换和亮度变换的数据增广方法对训练集进行扩充。为了在数据增强的前提下减少数据倾斜带来的副作用，本文分别对样本多的类别和样本少的类别采取不同的增广策略。当研究区内的木荷、马尾松、杉木的样本占比大于 70% 时，将以上数据增强方法随机排列组合，使此类样本的数量扩充到原来样本的 5 倍；反之也使用以上数据增强方法的随机排列组合，将该类别样本的数量扩充到原来样本的 10 倍。最终的样本集包括训练集 13 920 张、测试集 639 张。

2 研究方法

本文使用无人机可见光影像开展森林树种分类研究，技术路线如图 3 所示，主要包括数据获取、数据处理、模型构建与模型应用评价 4 个部分。在数据获取阶段，通过 Pix4Dmapper 软件将预处理后的无人机可见光影像进行拼接，拼接后生成 DOM 和 DSM；在数据处理阶段，首先综合利用三维信息和光谱特征获取试验区的森林冠层区域，其次结合森林小班和现有知识，基于 A、B、C、D 试验区构建树种的样本集，按照固定比例（8:2）将样本集划分为训练集和测试集，通过旋转、镜像、动态模糊等方法对训练集进行样本增强；在模型构建阶段，建立了基于 ResNet14* 的 UNet 树种分类模型，采用基于交叉熵和 Dice 系数的联合损失函数用于加快模型收敛，并和 4 个不同编码器（VGG16、ResNet18、ResNet34、ResNet50）进行精度对比；最后分析 Self-training 和 Mean teacher 两种不同的半监督学习方法在无人机影像中的模型泛化能力。

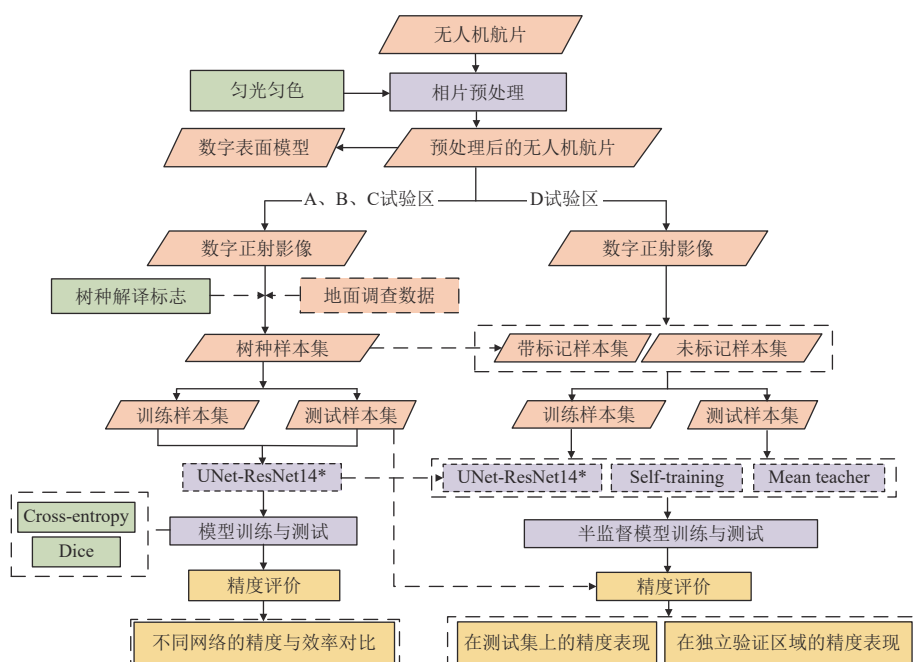


图 3 树种分类技术路线

Fig.3 Technical route of tree species classification

2.1 ResNet14*树种分类模型

2.1.1 ResNet18 模型改进

UNet 是由 RONNEBERGER 等^[15]提出的经典的语义分割模型,它采用与全卷积网络(fully convolutional networks, FCN)同样的 Encoder-Decoder 结构,左侧使用卷积方法提取上下文信息,右侧使用上采样方法定位语义信息以及还原至原始的影像分辨率。与 FCN 相比,UNet 采用了叠加的跳跃连接的方式,将低层语义信息与深层语义信息拼接在一起提高分割精度。

本文树种分类模型采用 UNet 网络作为树种分类的主体架构,设计的 UNet 网络如图 4 所示,主要包括下采样模块、上采样模块和跳跃连接 3 个部分。模型编码器是模型的下采样模块的主要部分,模型编码器常使用广泛应用的卷积神经网络(convolutional neural networks, CNN)模型,本文使用的是改进后的 ResNet18 网络即 ResNet14*网络来提取树种特征,与 ResNet18 相比,ResNet14*具有以下几个特点:ResNet14*去除了 ResNet18 中的 layer4 部分,即最后一个下采样的级联块,从而能保留稍高的空间信息;在 ResNet14*的 layer2、layer3 部分输出前增加一个最大值池化层,从而在保留原特征的同时减少神经网络训练的参数。

在 ResNet14*的每个级联块都由一个最大值池化层

和两个残差单元组成,其中 layer1 级联块的最大值池化层在级联块的前面,与 ResNet18 网络相对应。在本文的每个残差单元,用于下采样的卷积核采用 3×3 的卷积尺寸,用于缩放输入特征的卷积核采用 1×1 的卷积尺寸来保持特征信息。当特征图陆续进入到 3 个级联块(layer1、layer2、layer3)中时,每经过一个级联块,特征图像的大小减小一半,通道数量增加一倍;经过整个 ResNet14*主干网络后,图像的尺寸经过 4 次下采样,第一次下采样时,由进入 ResNet14*时的第一个卷积单元来实现下采样,4 次下采样后图像尺寸缩小到原来的 $1/4$,图像通道数扩展为 256。

在特征上采样的过程中采用的是线性插值方法,与其他上采样方法相比,线性插值方法具有速度快、占用空间小的优点,在一定程度上能够提高模型运算的效率。在上采样阶段,需要将上采样后的特征图像与下采样期间的特征图像进行拼接,然而 2 个特征图像之间的语义特征差别较大,直接连接可能加大模型训练的难度。因此,本研究将二者连接之前,先将下采样过程中的图像经过一个卷积核为 1×1 的卷积单元,通过卷积单元后的特征尺寸和通道数都不变,再对 2 个特征图像进行拼接,以此来减少前后的语义差异。

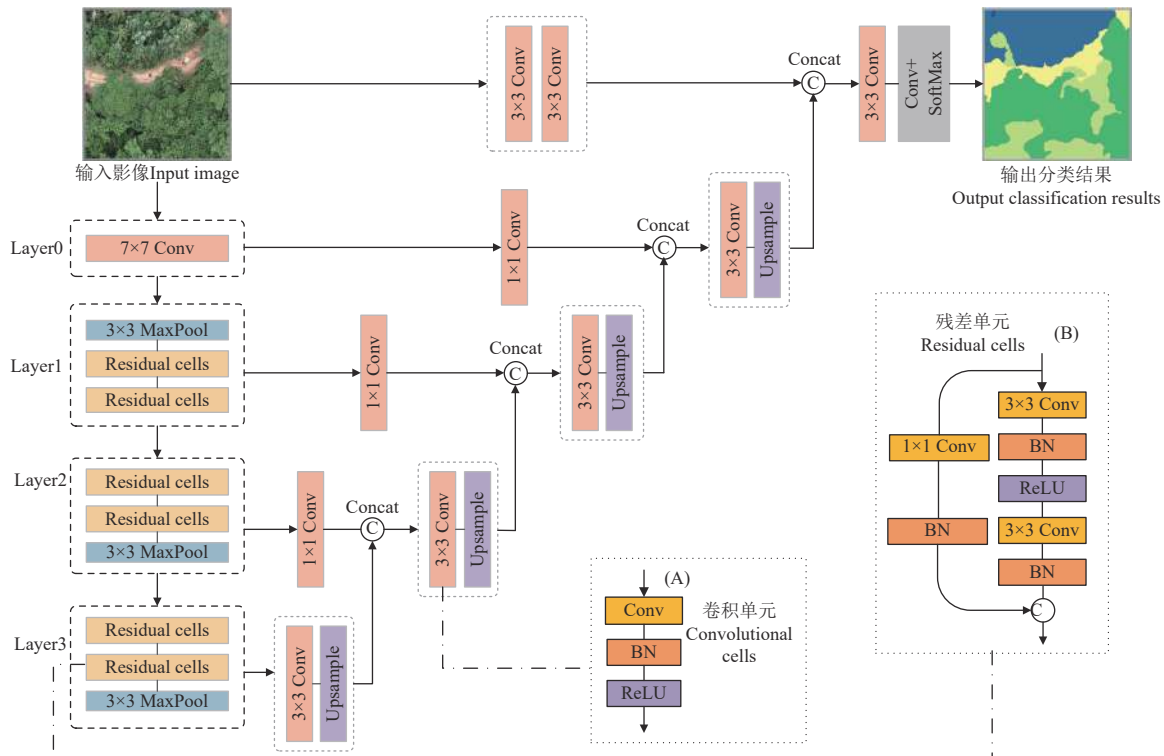


图 4 基于 ResNet14*树种分类模型
Fig.4 UNet architecture based on ResNet14*

在跳跃连接时,下采样的特征图和与之尺寸对应的上采样特征图在语义上存在着一定的差距,为此本研究将经过跳跃连接前的下采样特征图经过一个卷积单元来缓解这种语义差距;在模型的最后阶段,为了将原始图像的特征加入到特征图当中,本研究将原始图像投入 2 个卷积核为 3×3 的卷积单元,将上采样的特征图像与卷积结果拼接,并将拼接后的结果通过一个简单的卷积层和 SoftMax 层输出为类别信息。

2.1.2 ResNet14*模型的损失函数

在一般的人工林林班中,常见树种往往只有几种,其他树种只在林班内零星存在,因此在构建样本时往往会出现明显的长尾效应。在模型训练阶段,由于部分树种的样本量大,这些树种很容易成为影响损失函数大小的主要因素,从而会影响模型对占比较少的树种的分类效果。为此,本文采用交叉熵和 Dice 系数联合损失函数,用于解决树种样本数量不均衡造成的模型性能衰退的问题。

本文的联合损失函数为交叉熵和 Dice 系数损失函数指标的加权和, 见式 (1)。

$$L = \omega L_{CE} + (1 - \omega) L_{DL} \quad (1)$$

式中 L_{CE} 是交叉熵损失函数; L_{DL} 是 Dice 损失函数; ω 、 $(1 - \omega)$ 分别是 L_{CE} 、 L_{DL} 的权重。

交叉熵损失函数是最常见的用于衡量预测值与真值之间差异的损失函数, 交叉熵也被广泛应用在多种二分类和多分类问题中。该损失函数见式 (2)。

$$L_{CE} = - \sum_{i \in N} \sum_{l \in L} [y_i^{(l)} \lg \hat{y}_i^{(l)} + (1 - y_i^{(l)}) \lg (1 - \hat{y}_i^{(l)})] \quad (2)$$

式中 N 是所有像元的集合; L 是所有类别的集合; $y_i^{(l)}$ 是像元 i 在第 l 类别上的 one-hot 编码 (0 或 1); $\hat{y}_i^{(l)}$ 表示像元 i 在第 l 类别上的预测概率, $\hat{y}_i^{(l)} \in [0, 1]$ 。

Dice 系数通过比较预测和标签的交集来衡量模型的拟合效果。Dice 损失函数由 Dice 系数推导而来。该损失函数见式 (3)。

$$L_{DL} = 1 - 2 \frac{\sum_{l \in L} \sum_{i \in N} y_i^{(l)} \hat{y}_i^{(l)} + \varepsilon}{\sum_{l \in L} \sum_{i \in N} (y_i^{(l)} + \hat{y}_i^{(l)}) + \varepsilon} \quad (3)$$

式中 L 是所有树种类别的集合; $y_i^{(l)}$ 是像元 i 在第 l 类别上的 one-hot 编码 (0 或 1); $\hat{y}_i^{(l)}$ 表示像元 i 在第 l 类别上的预测概率, 其中 ε 是一个很小的常数, 本文中 $\varepsilon = 1 \times 10^{-8}$ 。

2.2 基于半监督学习的森林树种分类模型构建

半监督学习是指在有限的标记样本的情况下结合标记和未标记数据, 通过大量的未标记数据来增强分类器的泛化能力。半监督学习方法融合了有标签和未标签的数据, 从而实现了更好的模型性能。常用的半监督学习方法可以分为基于伪标签的方法和基于一致性正则化的方法。基于伪标签的方法是半监督学习的一种常见方法, 它使用有标记的数据训练模型后, 利用训练出的模型对未标记的数据进行预测, 并将分类置信度高的样本作为伪标签, 伪标签加入有标记数据中并用于下一轮训练, 以此来提高模型性能。常见的方法有自训练方法 (Self-training) [16-18] 和生成对抗网络 [19-21] 等。基于一致性正则化的方法也是常用的半监督方法 [22-24]。一致性正则化是在对同一个样本应用不同的扰动 (噪声) 时, 确保其预测结果不会发生显著变化。由于这种方法一般基于语义分割模型输出的预测向量, 不需要具体的标签, 因此能较好地适用于半监督学习领域。

2.2.1 基于 Self-training 的树种分类模型

在语义分割任务中, Self-training 方法可以通过预测

出的伪标签来扩充有标记的数据集, 以增强模型的泛化能力。Self-training 方法的主要流程如下 (图 5), 首先使用已有的带标记样本集来训练一个初始模型, 并用该模型对未标记的数据进行预测, 从而得到未标记数据的伪标签; 然后将这些带有伪标签的未标记数据加入到有标记数据集中, 重新训练模型。这个过程可以重复多次, 每次使用模型生成新的伪标签并将其加入到有标记的数据集中, 直到模型收敛或者达到预设的迭代次数为止。Self-training 方法的一个缺点是在伪标签错误的情况下可能会产生累积错误, 进而导致模型性能下降。因此, 为了避免标记错误的累积, 可以通过设置置信度阈值来筛选带有伪标签的未标记数据。

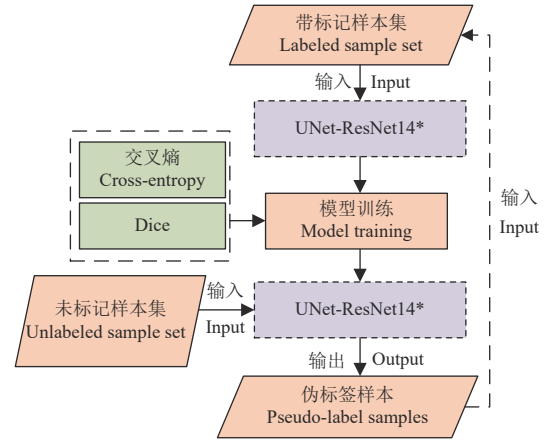


图 5 Self-training 方法框架结构

Fig.5 Method frame structure of Self-training

2.2.2 基于 Mean teacher 的树种分类模型

Mean teacher 是一种基于一致性正则化的半监督学习方法 [25], 本文主要基于 Temporal Ensembling 模型做出了一些改进。Mean teacher 的基本思路是在有标记的样本和无标记的样本上训练 2 个相同的神经网络, 一个作为 Teacher 模型, 另一个作为 Student 模型。其中 Student 模型的权重在每一批次后都会更新, 而 Teacher 模型的权重则是通过指数平均的方式更新, 未标记样本在 2 个模型上得到 2 个预测值, 通过这种方式可以缓解 Temporal Ensembling 在训练时的振荡和噪声影响, 批次每更新一次参数就更新一次, 从而实时更新模型参数和提高模型性能。本文提出的 Mean teacher 模型结构如图 6 所示, 其中 Student 模型和 Teacher 模型均是上文提出的 UNet-ResNet14* 模型。

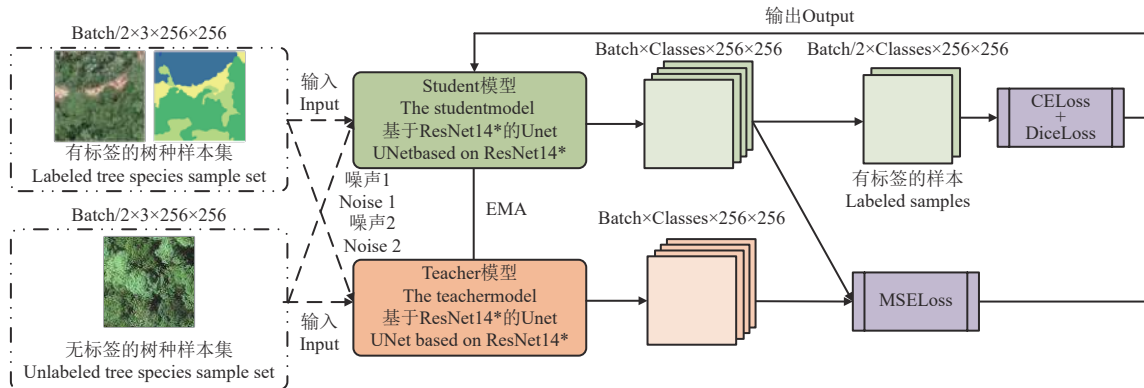


图 6 Mean teacher 方法框架结构

Fig.6 Method frame structure of Mean teacher

2.3 模型的训练与测试

本文所有实验均在以下配置的服务器上运行：操作系统为 Ubuntu18.04，CPU 为 12 核的 Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50 GHz，显卡为 RTX3090(24G)，运行内存为 24GB，编程语言为 Python 3.8，深度学习框架为 PyTorch 1.8.1。在模型训练时，本文使用 Adam 优化器作为模型损失函数的优化器，分别设置初始学习率为 0.01、0.001 和 0.000 1，同时采用多步学习率衰减策略用于缩小学习率，gamma 值为 0.5。经过实验，发现以 Adam 作为优化器的情况下初始学习率设置为 0.001 训练出的模型分类精度最高，因此后续实验均选择 0.001 作为初始学习率。此外本文还采取以下的训练策略：采用迁移学习的方法，所有模型的编码器均使用 ImageNet 数据集的预训练权重来进行参数初始化，以加快模型收敛速率；在模型训练的每次迭代结束后，都将测试样本集纳入模型进行预测；为了防止模型退化，在测试集连续 3 次迭代时精度都在下降，模型训练停止；所有模型的批处理大小均设置为 30，所有的模型训练均迭代 50 次，选择出在测试集上精度最高的模型参数。

2.4 模型的评价指标

为评估模型的综合性能，本研究从总体精度（overall

accuracy, OA）、准确率（precision）、召回率（recall）、Kappa 系数、类别像素精度（pixel accuracy, PA）预测时间（forecast time）等多个指标来评估模型在树种样本集中的表现。

3 结果与分析

3.1 UNet-ResNet14*模型的分类结果

为了证明基于 ResNet14*的 UNet 网络的性能优势，本文构建了基于不同编码器的 UNet 网络，并从模型精度、模型参数以及预测效率多个角度对模型进行对比，每个编码器均保持了原始网络的结构。本文对比了上文中的 5 个 UNet 模型，其中每个网络联合损失函数均设置为 0.5，每个网络的各项统计指标如表 2 所示。从编码器结构上可以发现，以 ResNet 为主干的模型能够得到比以 VGG16 为主干的模型更高的总体精度。在总体精度上，ResNet14*网络的总体精度达到 91.15%，Kappa 系数达到 0.827，与其他 ResNet 模型的精度相差在 1% 内，以 VGG16 为主干的模型的总体精度仅达到 89.92%，比以 ResNet 为主干的树种分类模型的总体精度都低；在参数量上，以 ResNet14*为主干的模型的参数量也比以 VGG16 为主干的模型的参数量小。

表 2 不同骨干网络在测试集上的效率表现和树种分类精度

Table 2 Efficiency performance and tree species classification accuracy of different backbone networks on the test set												
骨干网络 Backbone networks	总体精度 Overall accuracy/%	精确率 Precision	召回率 Recall	Kappa	参数大小 Parameter size/MB	预测时间 Forecast time/s	树种分类精度 Tree species classification accuracy/%					
							木荷 Gugertree	马尾松 Masson pine	杉木 Chinese fir	竹类 Bamboo	福建柏 Fujian cypress	其他阔叶树 Other broad-leaved trees
ResNet14*	91.15	0.732	0.784	0.827	19.14	35.35	90.00	85.29	83.54	86.71	86.55	20.29
ResNet18	91.19	0.750	0.770	0.829	67.92	36.87	88.04	88.87	78.93	88.01	90.75	34.41
ResNet34	91.45	0.763	0.711	0.841	106.48	38.29	88.84	84.29	81.22	73.23	91.47	36.86
ResNet50	91.36	0.755	0.717	0.834	180.48	39.74	88.39	87.87	79.36	75.83	91.29	37.16
VGG16	89.92	0.699	0.700	0.805	88.60	35.82	83.50	90.93	76.11	70.88	71.84	20.68

事实证明，针对于本研究的数据集，基于 ResNet 构建的树种分类模型能够得到比基于 VGG16 构建的树种分类模型更好的分类效果。在树种类别分类精度上，5 类树种（木荷、马尾松、杉木、竹类、福建柏）分别在 ResNet14*、VGG16、ResNet14*、ResNet18、ResNet34 五个模型上分类精度最高，其中样本量充足的木荷、马尾松、杉木三个类别在不同深度的 ResNet 模型中精度都差距不大，ResNet14*在木荷和杉木两个类别中表现较为突出；竹类在 ResNet14*和 ResNet18 两个模型中精度较高，在 ResNet34 和 ResNet50 两个模型中精度较低；VGG16 模型除了在马

尾松类别上表现出较好的精度以外，在其他所有类别中精度都较差。综上所述，ResNet14*相比于其他模型具有一定优势。

3.2 基于半监督学习模型的分类结果

3.2.1 在测试集上的精度表现

本文将以上 2 种半监督方法得到的模型在原有测试集上进行预测，精度如表 3 所示。与原始模型相比，Self-training 方法在原有测试集上的总体指标差别不大，总体精度达到 91.08%，Kappa 系数甚至比原有模型还高一些，但是 Mean teacher 模型的总体指标相对较差，各项指标均比原有模型低。

表 3 不同模型在原始测试集中的效率表现和树种分类精度

Table 3 Efficiency performance and tree species classification accuracy of different models in the original test set										
模型 Model	总体精度 Overall accuracy/%	精确率 Precision	召回率 Recall	Kappa	树种分类精度 Tree species classification accuracy/%					
					木荷 Gugertree	马尾松 Masson pine	杉木 Chinese fir	竹类 Bamboo	福建柏 Fujian cypress	其他阔叶树 Other broad-leaved trees
UNet-ResNet14*	91.15	0.732	0.784	0.827	90.00	85.29	83.54	86.71	86.55	20.29
Self-training	91.08	0.732	0.748	0.828	92.86	91.64	93.33	84.55	87.09	12.08
+UNet-ResNet14*	89.92	0.679	0.698	0.785	89.74	90.06	94.68	69.60	73.90	4.70

从各树种的类别精度上看，经过 Self-training 之后的模型在木荷、马尾松、杉木 3 个样本充足的类别中

精度均有所提升，说明 Self-training 方法创造的伪标签可以带来更强的模型泛化能力。Mean teacher 方法在

各个类别的分类精度表现出大范围的波动，在马尾松和杉木类别中各提升了 5% 和 10% 的分类精度，木荷类别的精度略微下降，而对于竹类、福建柏和其他阔叶树三个类别精度均大幅度降低，其中竹类精度下降了 17% 左右，其他阔叶树的精度更是下降到个位数，因此本研究推测样本不平衡对于 Self-training 方法的精度影响不大，在原始样本集和未标记样本集的分类比例不一致的情况下也能带来较好的精度提升，而 Mean teacher 方法容易受到原始样本集和未标记样本集的样本量和样本比例的影响，这也可能与 Mean teacher 方法超参数的设置有关。

表 4 不同模型在验证区 D 中的效率表现和树种分类精度

模型 Model	总体精度 Overall accuracy/%	精确率 Precision	召回率 Recall	Kappa	树种分类精度 Tree species classification accuracy/%					
					木荷	马尾松	杉木	竹类	福建柏	其他阔叶树
					Gugertree	Masson pine	Chinese fir	Bamboo	Fujian cypress	Other broad-leaved trees
UNet-ResNet14*	85.42	0.619	0.554	0.766	80.68	93.40	82.30	24.22	24.61	95.47
Self-training										
+UNet-ResNet14*	88.50	0.633	0.600	0.811	80.78	92.43	87.82	23.08	31.39	95.41
Mean teacher										
+UNet-ResNet14*	73.56	0.609	0.499	0.615	72.80	92.12	63.51	30.69	29.16	94.67

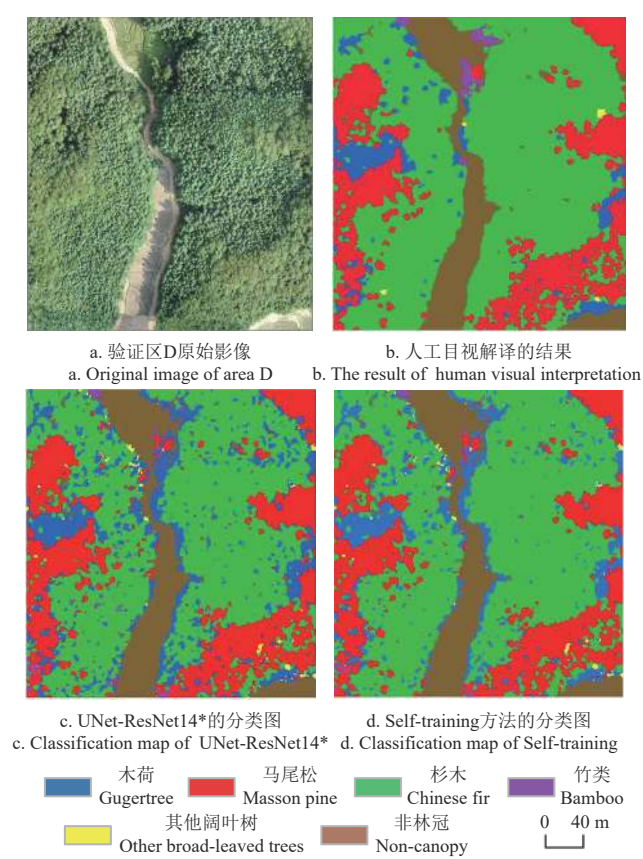


图 7 Self-training 与 UNet-ResNet14*的树种分类图

Fig.7 Map of tree species classification of Self-training and UNet-ResNet14*

在各个类别的精度表现上，样本量充足的木荷、马尾松、杉木即使使用原始模型，各类别的精度也都在 80% 以上，Self-training 方法在杉木类别上精度表现更好，从验证区 D 的无人机影像上看，50% 以上的区域都是杉木，在构建伪标签时这种大面积分布的杉木也为 Self-training 带来了便利条件，相较于样本数量较多的以上 3

3. 2. 2 在独立验证区域上的精度表现

本实验将以上两种半监督方法和原有模型在独立验证区域（试验区 D）上进行对比，精度表现如表 4 所示，分类结果如图 7 所示。其中基于伪标签的 Self-training 方法的总体精度相比于原有模型精度有所提升，达到了 88.50%，而基于一致性损失的 Mean teacher 模型的总体精度相比于原有模型精度大幅度降低，总体精度降低到 80% 以下；Self-training 方法的在精确率、召回率和 Kappa 系数上的表现都比原始模型好，而 Mean teacher 方法的表现相较于原始模型有了大幅度的模型退化。

个类别，竹类在影像上仅有小片区域零星分布，甚至不足试验区总面积的 1/20，而竹类的冠层特征也复杂多变，因此在 3 种方法中精度都相对较低，其他阔叶树在验证区 D 中零星分布，精度表现也较差。

综合 3 种方法的精度表现，Self-training 方法精度在原始模型的基础上提升了 3% 左右，Mean teacher 方法则降低了 12% 左右。由实验可知，从样本集的数量上看，原始训练集和未标记的样本集比例为 5:3，在常见的 Mean teacher 方法的应用中，未标记样本集往往要比原始样本集多，这可能是 Mean teacher 方法在本数据集中精度降低的一个原因；此外 Mean teacher 模型可能将添加噪声前后的未标记的样本均预测为其他类别，从而使模型的误差越来越大。而 Self-training 方法依靠置信概率来判断样本是否可以用于后续分类，同时本研究的原始模型训练的精度就比较高，未标记样本也可以得到较为准确的伪标签信息，综上所述，Self-training 方法比 Mean teacher 方法更适用于检测在其他试验区的模型泛化能力。

4 讨 论

本研究基于卷积神经网络实现了人工林优势树种的自动分类，分析 Self-training 和 Mean teacher 两种不同的半监督学习方法在无人机影像中的模型泛化能力，证明且分析了无人机可见光影像在福建省常见人工林树种分类中的可行性。以 UNet-ResNet14*为主网络的 Self-training 分类模型，对 ResNet18 模型进行剪裁来兼顾模型的效率和精度，最终在保持原来分类精度的条件下将模型参数大小降低到原来参数量的 1/3。大部分的研究在设计轻量化模型时，均是基于现有的主干网络和语义分割架构剪裁^[26-27]，有的研究也会使用知识蒸馏、模型剪枝等方法进一步减少模型的参数大小与计算量^[28]。如刘慧

等^[26]使用轻量化 MobileNet V2 网络替代原有的 Xception 网络以减少网络参数,模型参数量较改进前压缩了 88.67%;LAN 等^[28]对 Swin Transformer 的多层堆叠结构进行剪枝,提出了一种轻量的 Swin-T YOLOX 模型用于监测果树病虫害区域,模型参数大小仅为 29.0 M。在现有的半监督学习研究中,有的研究会融合监督学习和半监督学习来提高模型的学习能力^[29],有的研究也会使用混合数据增强和注意力机制来提高模型的精度与性能^[30]。如王瑞琪等^[29]整合监督损失优化半监督损失,使模型学习更全面的信息;吕佳等^[30]在 YOLOv5 的骨干网络中加入条带注意力模块提高模型的识别能力。后续研究还可以综合多个方法进一步减少本文模型的复杂度,并优化样本数据增强方法,以减少模型参数量和优化模型的学习能力。

总体而言,本研究的树种分类方法具有一定的精度效果,但仍然存在一些局限性。首先本文的试验区主要以木荷、马尾松、杉木、竹类、福建柏为主要研究的树种,后续可以进一步对树种类别进行扩充;(2)本研究设置为固定的分辨率,后续可以研究在多个分辨率上都能适用的树种分类模型;(3)本研究的试验区树龄均为 20~30 年,树龄组成较为单一,模型在幼林地上的普适性还需要进一步的研究。(4)本研究中没有按照地形或者郁闭度划分森林中的区域,后续可以在此基础上开展研究,进而评价模型在不同场景中的精度表现。(5)本研究每年仅获取一次无人机影像进行研究,未考虑树种的生长状态在一年中存在差异,后续可以分析模型在不同时间上的迁移性。

5 结 论

本文基于卷积神经网络提出了一种以 ResNet14*为模型主干的半监督方法的树种分类框架,该模型基于广泛使用的无人机可见光影像作为数据源,对高郁闭度条件下的福建省人工林实现了树种的自动分类识别。主要结论如下:

1) 实验构建了以 ResNet14*为主干模型的 UNet 网络作为树种分类网络模型,并与 VGG16、ResNet18、ResNet34、ResNet50 网络进行了对比分析。结果表明,ResNet14*网络的总体精度达到 91.15%,Kappa 系数达到 0.827,与其他 ResNet 模型的精度相差在 1% 内,同时模型参数量更小且预测时间最短,兼顾了树种分类的精度和效率。

2) 本文在独立验证区域上进行精度评价,试验区面积占比较多的木荷、马尾松、杉木三类树种的分类精度均在 80% 以上,一定程度上满足了试验区的树种制图精度需求。

3) 分别实现了以 UNet-ResNet14*为主网络的 Self-training 和 Mean teacher 两种半监督学习方法。实验发现,Self-training 方法在测试集上的总体精度达到 91.08%,比原始模型略低,但是在样本充足的木荷、马尾松、杉木类别上得到了比原始模型更高的类别精度。将 2 个半

监督模型在试验区 D 上进行对比,其中基于伪标签的 Self-training 方法的总体精度相比于原有模型精度有所提升,达到了 88.50%;而基于一致性损失的 Mean teacher 模型的总体精度相比于原有模型精度大幅度降低。本研究认为,在大面积树种制图时可以采用 Self-training 方法模型来实现林区优势树种的快速有效制图。

[参 考 文 献]

- [1] 付波霖,孙军,李雨阳,等.基于多光谱影像和机器学习算法的红树林树种 LAI 估算[J].农业工程学报,2022,38(7): 218-228.
FU Bolin, SUN Jun, LI Yuyang, et al. Mangrove LAI estimation based on remote sensing images and machine learning algorithms[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(7): 218-228. (in Chinese with English abstract)
- [2] 郑迪,沈国春,王舶鉴,等.基于无人机高光谱影像和深度学习算法的长白山针阔混交林优势树种分类[J].生态学杂志,2022,41(5): 1024-1032.
ZHENG Di, SHEN Guochun, WANG Bojian, et al. Classification of dominant species in coniferous and broad-leaved mixed forest on Changbai Mountain based on UAV-based hyperspectral image and deep learning algorithm[J]. Chinese Journal of Ecology, 2022, 41(5): 1024-1032. (in Chinese with English abstract)
- [3] 徐岩,张聪伶,降瑞娇,等.无人机高光谱影像与冠层树种多样性监测[J].生物多样性,2021,29(5): 647-660.
XU Yan, ZHANG Congling, JIANG Ruijiao, et al. UAV-based hyperspectral images and monitoring of canopy tree diversity[J]. Biodiversity Science, 2021, 29(5): 647-660. (in Chinese with English abstract)
- [4] LIU M, HAN Z, CHEN Y, et al. Tree species classification of LiDAR data based on 3D deep learning[J]. Measurement, 2021, 177: 109301.
- [5] 李佳柠,李明泽,全迎,等.无人机激光雷达与高光谱数据协同的帽儿山地区树种分类[J].东北林业大学学报,2022,50(6): 63-69;116-116. (in Chinese with English abstract)
LI Jianing, LI Mingze, QUAN Ying, et al. Tree species classification using UAV-LiDAR with hyperspectral data[J]. Journal of Northeast Forestry University, 2022, 50(6): 63-69; (in Chinese with English abstract)
- [6] 李华玉,张超,陈巧,等.基于无人机可见光和 LiDAR 数据的单木树种识别[J].西南林业大学学报(自然科学),2021,41(5): 105-113.
LI Huayu, ZHANG Chao, CHEN Qiao, et al. Single tree species identification based on visible light and LiDAR data of UAV[J]. Journal of Southwest Forestry University(Natural Sciences), 2021, 41(5): 105-113. (in Chinese with English abstract)
- [7] ODDI L, CREMONESE E, ASCARI L, et al. Using UAV

- imagery to detect and map woody species encroachment in a subalpine grassland: advantages and limits[J]. *Remote Sensing*, 2021, 13(7): 1239.
- [8] HAO Z, LIN L, POST C J, et al. Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN)[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 178: 112-123.
- [9] FENG X, LI P. A tree species mapping method from UAV images over urban area using similarity in tree-crown object histograms[J]. *Remote Sensing*, 2019, 11(17): 1982.
- [10] 蒲涛, 王妮, 龚育红, 等. 基于无人机影像的树种分割实践[J]. *林业资源管理*, 2023(1): 115-126.
- PU Tao, WANG Ni, GONG Yuhong, et al. Tree species segmentation practice based on UAV imagery[J]. *Forest Resources Management*, 2023(1): 115-126. (in Chinese with English abstract)
- [11] 罗仙仙, 许松芽, 陈桂莲, 等. 基于预训练模型与无人机可见光影像的树种识别[J]. *计算机系统应用*, 2022, 31(7): 386-391.
- LUO Xianxian, XU Songya, CHEN Guilian, et al. Tree species identification by visible image from pre-trained model and UAV[J]. *Computer Systems & Applications*, 2022, 31(7): 386-391. (in Chinese with English abstract)
- [12] WANG X, WANG Y, ZHOU C, et al. Urban forest monitoring based on multiple features at the single tree scale by UAV[J]. *Urban Forestry & Urban Greening*, 2021, 58: 126958.
- [13] NATESAN S, ARMENAKIS C, VEPAKOMMA U. Individual tree species identification using Dense Convolutional Network (DenseNet) on multitemporal RGB images from UAV[J]. *Journal of Unmanned Vehicle Systems*, 2020, 8(4): 310-333.
- [14] NEZAMI S, KHORAMSHAHI E, NEVALAINEN O, et al. Tree species classification of drone hyperspectral and RGB imagery with deep learning Convolutional Neural Networks[J]. *Remote Sensing*, 2020, 12(7): 1070.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. *MICCAI*, 2015, 3: 234-241.
- [16] ZHENG X, JIA J, CHEN J, et al. Hyperspectral image classification with imbalanced data based on semi-supervised learning[J]. *Applied Sciences*, 2022, 12(8): 3943.
- [17] GEUNHO K, NOWOOK P. Unsupervised domain adaptation with adversarial Self-Training for crop classification using remote sensing images[J]. *Remote Sensing*, 2022, 14(18): 4639-4639.
- [18] CUI Hongyan, WANG Gangkun, LI Yuanxin, et al. Self-training method based on GCN for semi-supervised short text classification[J]. *Information Sciences*, 2022, 611: 18-29.
- [19] 刘少鹏, 洪佳明, 梁杰鹏, 等. 面向医学图像分割的半监督条件生成对抗网络[J]. *软件学报*, 2020, 31(8): 2588-2602.
- LIU Shaopeng, HONG Jiaming, LIANG Jiepeng, et al. Medical image segmentation using semi-supervised conditional generative adversarial nets[J]. *Journal of Software*, 2020, 31(8): 2588-2602. (in Chinese with English abstract)
- [20] 李季兰. 基于深度对抗生成网络的彩超图像半监督分类研究[J]. *现代计算机*, 2020(30): 47-51.
- LI Jilan. Semi-supervised classification for ultrasound images based on GAN[J]. *Modern Computer*, 2020(30): 47-51. (in Chinese with English abstract)
- [21] HU Han, LIANG Xinrong, DING Yulin, et al. Semi-supervised adversarial recognition of refined window structures for inverse procedural façade modelling[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 192: 215-231.
- [22] 蔡雨霖, 肖佳仪, 余超然, 等. 基于 UANP-MT 的半监督菜心杂草分割方法[J]. *农业工程学报*, 2023, 39(11): 183-191.
- CAI Yulin, XIAO Jiayi, YU Chaoran, et al. UANP-MT based semi-supervised image segmentation method for identifying weeds in cabbage field[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(11): 183-191. (in Chinese with English abstract)
- [23] LI Lerenhan, DONG Yunlong, REN Wenqi, et al. Semi-supervised image dehazing[J]. *IEEE Transactions on Image Processing*, 2019, 29: 2766-2779.
- [24] OUALI Y, HUDELLOT C, TAMI M. Semi-supervised semantic segmentation with cross-consistency training[J]. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020: 12671-12681.
- [25] TARVAINEN A, VALPOLA H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 1196-1205.
- [26] 刘慧, 姜建滨, 沈跃, 等. 基于改进 DeepLab V3+ 的果园场景多类别分割方法[J]. *农业机械学报*, 2022, 53(11): 255-261.
- LIU Hui, JIANG Jianbin, SHEN Yue, et al. Multi-category segmentation of orchard scene based on improved DeepLab V3+ [J]. *Transactions of the Chinese Society for Agricultural Machinery (Transactions of the CSAE)*, 2022, 53(11): 255-261. (in Chinese with English abstract)
- [27] LIU J, XIANG J, JIN Y, et al. Boost precision agriculture with unmanned aerial vehicle remote sensing and edge intelligence: a survey[J]. *Remote Sensing*, 2021, 13(21): 4387.
- [28] LAN Y, LIN S, DU H, et al. Real-time UAV patrol technology in orchard based on the Swin-T YOLOX lightweight model[J]. *Remote Sensing*, 2022, 14(22): 5806.
- [29] 王瑞琪, 纪淑娟, 曹宁, 等. 基于一致性训练的半监督虚假招聘广告检测模型[J]. *计算机应用*, 2023, 43(9): 2932-2939.
- WANG Ruiqi, JI Shujuan, CAO Ning, et al. Semi-supervised fake job advertisement detection model based on consistency

- training[J]. Journal of Computer Applications, 2023, 43(9): 2932-2939. (in Chinese with English abstract)
- [30] 吕佳, 李帅军, 曾梦瑶, 等. 基于半监督 SPM-YOLOv5 的套袋柑橘检测算法[J]. 农业工程学报, 2022, 38(18): 204-211.

LYU Jia, LI Shuaijun, ZENG Mengyao, et al. Detecting bagged citrus using a semi-supervised SPM-YOLOv5[J]. Transactions of the Chinese Society for Agricultural Machinery (Transactions of the CSAE), 2022, 38(18): 204-211. (in Chinese with English abstract)

Classification of tree species based on UNet-ResNet14* semi-supervised learning using UAV images

CHEN Longwei¹, ZHOU Xiaocheng^{1*}, LI Chuanxin¹, LIN Huazhang², WANG Yongrong³, CUI Yonghong⁴

(1. Local Joint Engineering Research Center of Satellite Geospatial Information Technology & Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National, Fuzhou University, Fuzhou 350108, China; 2. Fujian Datian Taoyuan State-owned Forest Farm, Sanming 366199, China; 3. Fujian Zhangping Wuyi State-owned Forest Farm, Longyan 364400, China; 4. Fujian Forest Inventory and Planning Institute, Fuzhou 350003, China)

Abstract: Unmanned aerial vehicle (UAV) remote sensing has the promising potential for the precise and efficient classification and mapping of forest tree species. Deep learning also requires a large number of datasets for training, typically on manual annotation. In this study, the framework of forest tree species classification was proposed to fully utilize a large amount of unlabeled data and a small amount of annotated data using semi-supervised learning. A rapid and accurate classification was also achieved in the high-precision distribution of dominant tree species in forests. The experimental areas were taken as the complex mountainous forest environment in Fujian Province. The composition of tree species was then obtained in a rapid, effective, and cost-saving manner. Taking four experimental areas in Fuzhou, Longyan, and Sanming in Fujian Province as examples, the simplified classification was constructed in the UNet tree species (ResNet14 *) model with ResNet18 as the backbone. ResNet14 * was different from ResNet18: ResNet14 * was used to remove the layer4 part of ResNet18, i.e., the last downsampled cascaded block, which retained slightly higher spatial information; At the end of the layer2 and layer3 sections of ResNet14 *, a max pooling layer was added to reduce the training parameters of the neural network while retaining the original features. A joint loss function of cross entropy and Dice coefficient was used to optimize the model parameters. The generalization of Self-training and Mean teacher was evaluated on the classification models with semi-supervised learning using UVA images. The results show that the overall accuracy of the ResNet14 * network reached 91.15%, with a Kappa coefficient of 0.827, which was within 1% of the accuracy of the rest ResNet models. At the same time, a smaller number of parameters and the shortest prediction time were achieved to balance the accuracy and efficiency of tree species classification. The best prediction performance of ResNet14 * was achieved with the joint loss function weight of 0.5, indicating an overall accuracy of 91.15%. Therefore, the joint loss function weight of 0.5 was an optimal value for semi-supervised learning in this case. Self-training and Mean teacher semi-supervised learning were implemented with UNet (ResNet14 *) as the main network. The experiment showed that the overall accuracy of the Self-training on the test set reached 91.08%, slightly lower than the original. The higher category accuracy was also achieved in the categories of *Schima superba*, *Pinus massoniana*, and Chinese fir with sufficient samples. Furthermore, the overall accuracy of the self-training with pseudo labels was improved among two semi-supervised models in experimental area D, reaching 88.50% compared with the original; There was a significant decrease in the overall accuracy of the Mean teacher model with consistency loss. The total accuracy of the Mean teacher model was 88.56%, where the accuracy was 73.56% in the independent validation area. Accuracy evaluation was also conducted on an independent validation area. The classification accuracy of above 80% was found in the three types of tree species, namely *Schima superba*, *Pinus Massoniana*, and Chinese fir. A relatively large area was accounted for to meet the accuracy requirements of tree species mapping in the experimental area. Therefore, the semi-supervised learning of the Self-training model can be expected to rapidly obtain the composition of tree species in the experimental area.

Keywords: UAV; Remote sensing; Forest; Tree species; Visible light; Unet; ResNet; Semi-supervised learning