

# 基于预训练模型与无人机可见光影像的树种识别<sup>①</sup>



罗仙仙<sup>1,2</sup>, 许松芽<sup>3</sup>, 陈桂莲<sup>1,2</sup>, 严洪<sup>4</sup>, 万晓会<sup>4</sup>

<sup>1</sup>(泉州师范学院 数学与计算机科学学院, 泉州 362000)

<sup>2</sup>(福建省大数据管理新技术与知识工程重点实验室, 泉州 362000)

<sup>3</sup>(泉州师范学院 教育科学学院, 泉州 362000)

<sup>4</sup>(福建省林业调查规划院, 福州 350003)

通信作者: 许松芽, E-mail: 29974817@qq.com

**摘要:** 提高图像质量与利用新的图像分类方法是提高遥感图像树种识别精度两个突破口。本文基于 VGG16 的预训练模型与无人机可见光影像进行杉木、马尾松 2 个树种识别研究。利用大疆精灵 4RTK 无人机, 搭载 FC6310R 相机, 采集南平市和三明市的杉木和马尾松人工纯林彩色图像。通过图像预处理、标注、裁剪和增强等环节构建 UAVTree2k 和 UAVTree20k 两个数据集。基于 UAVTree2k 数据集和 VGG16 模型在 ImageNet 数据集的预训练模型, 重新训练 3 个全连接层和 Sigmoid 层, 研究探讨不同迭代次数、不同批次大小、不同训练集和测试集划分比例对识别精度的影响。研究表明, 当迭代次数为 40、批次大小为 16、训练集和测试集为 6:4 时, 模型识别效果最好, 测试精度达到 98.63%; 小样本下, 基于 VGG16 的预训练模型具有良好的特征学习能力。

**关键词:** 迁移学习; 卷积神经网络; 无人机; 树种识别

引用格式: 罗仙仙, 许松芽, 陈桂莲, 严洪, 万晓会. 基于预训练模型与无人机可见光影像的树种识别. 计算机系统应用, 2022, 31(7): 386-391. <http://www.c-s-a.org.cn/1003-3254/8618.html>

## Tree Species Identification by Visible Image from Pre-trained Model and UAV

LUO Xian-Xian<sup>1,2</sup>, XU Song-Ya<sup>3</sup>, CHEN Gui-Lian<sup>1,2</sup>, YAN Hong<sup>4</sup>, WAN Xiao-Hui<sup>4</sup>

<sup>1</sup>(School of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China)

<sup>2</sup>(Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou 362000, China)

<sup>3</sup>(School of Educational Science, Quanzhou Normal University, Quanzhou 362000, China)

<sup>4</sup>(Fujian Forest Inventory and Planning Institute, Fuzhou 350003, China)

**Abstract:** Enhancing image quality and adopting new image classification methods are two breakthrough points to improve the accuracy of tree species identification by remote sensing images. The research focuses on the identification of Chinese fir and Masson pine by the pre-trained model of VGG16 and unmanned aerial vehicle (UAV) visible images. The DJI Phantom 4RTK UAV with an FC6310R camera is used to collect color images of artificial pure forests of Chinese fir and Masson pine in Nanping and Sanming cities. Then, two datasets UAVTree2k and UAVTree20k are constructed through image preprocessing, annotation, cropping, and enhancement. Furthermore, three full connection layers and Sigmoid layer are trained by the UAVTree2K dataset and the pre-trained model of VGG16 on the ImageNet dataset to investigate the effects of the number of iterations, batch size, partition ratios of the training set and the test set on identification accuracy. The results show that when the number of iterations is 40, the batch size is 16, and the ratio between the training set and the test set is 6:4, the identification effect of the model is best, and the test accuracy reaches 98.63%. Meanwhile, the VGG16-based pre-trained model has a good feature learning ability for a small sample size.

**Key words:** transfer learning; convolutional neural network; unmanned aerial vehicle; tree species identification

① 基金项目: 福建省自然科学基金 (2020J01785); 泉州市科技计划 (2021N180S); 泉州师范学院大学生创新创业训练计划 (202110399021)

收稿时间: 2021-09-15; 修改时间: 2021-10-14; 采用时间: 2021-12-14; csa 在线出版时间: 2022-05-31

人工智能是未来的第一生产力,人工智能技术与方法在我国林业中的应用初见端倪.遥感图像的树种识别始终是国际关注的学术前沿<sup>[1]</sup>.无人机低空航拍受天气影响小,作业方式灵活,获取的影像比例尺大,精度高,为森林资源遥感调查研究提供了新的技术手段<sup>[2]</sup>.然而,传统人工目视解译方法和遥感图像处理方法,并不能够对无人机遥感图像进行自动识别,而且精度难以保证.

深度学习是人工智能研究中重要方法,2006年加拿大多伦多大学 Hinton 等提出<sup>[3]</sup>,能够模拟人的神经系统,可以从大量数据中自动提取特征.深度学习方法用于无人机树种识别的研究刚刚起步.由于深度学习方法依赖于大量的训练数据与测试数据,并且数据标注费时费力,一定程度限制深度学习在遥感图像树种识别研究,因此,少样本的迁移学习方法是当前研究热点.目前,尚没有开放的无人机树种识别数据集<sup>[1]</sup>,武红敢等人指出林业遥感专用数据库的必要性和紧迫性,并提出建设思路和原则<sup>[4]</sup>.

迁移概念最早由心理学家提出,一般认为,迁移是指一种学习对于另一种学习的影响<sup>[5]</sup>.在人工智能和机器学习领域,迁移学习是一种学习的思想和模式,可以利用数据、任务或模型之间的相似性,将在旧领域学习过的模型和知识应用于新的领域.基于模型的迁移学习方法(预训练方法)是指从源域和目标域中找到它们之间共享的参数信息,以实现迁移方法<sup>[6]</sup>.预训练模型在大数据集(如 ImageNet<sup>[7]</sup>)上训练得到一个具有强泛化能力的模型,作为后续任务的基准模型,通过少量有标注的数据微调训练可以取得较好识别效果.

近两年,林业遥感领域的研究学者把迁移学习应用于无人机影像的树种识别.王莉影等<sup>[8]</sup>较早应用迁移学习思想,基于 Inception-v3 模型在 ImageNet 数据集上进行预训练,然后利用有限样本重新训练一个全连接层,对东北林业大学实验林场只进行针叶林、阔叶林分类研究,总体分类精度为 98.4%, Kappa 系数为 0.987.滕文秀等<sup>[9]</sup>基于迁移学习思想,对 4 种常规的卷积神经网络方法在 ImageNet 数据集上进行预训练后,增加全局平均池化层、1 个全连接层和 1 个 Softmax 层,对 4 个树种进行分类,研究表明 VGG16 作为树种分类模型的预训练模型是好的. Natesan 等<sup>[10]</sup>基于 ResNet50 模型在 ImageNet 数据集上进行预训练后,增加最大池化层和 4 个全连接层,基于 Keras 框架进行

3 类树种分类,总体精度达 80%.总之,现有研究没有对数据集制作详细介绍,数据集尚未共享一定程度影响此领域迅速发展.

论文利用 TensorFlow 框架,基于 VGG16 在 ImageNet 数据集上预训练模型,进行无人机可见光影像树种识别研究.利用大疆精灵 4RTK 无人机,搭载 FC6310R 相机,采集南平市和三明市的杉木和马尾松人工纯林彩色图像.通过图像预处理、标注、裁剪和增强等环节构建两个数据集 UAVTree2k 和 UAVTree20k.基于 UAVTree2k 数据集和 VGG16 模型在 ImageNet 数据集的预训练模型,重新定义 3 个全连接层,输出层设置成 2,进行小样本的杉木和马尾松二分类研究;验证小样本的预训练模型的有效性,以期复杂林分环境下的树种分类与制图提供参考.

## 1 数据集的构建

### 1.1 无人机外业数据采集

杉木与马尾松数据采集,均采用无人机大疆精灵 4RTK,搭载机型为 FC6310R,相机像素为 2 000 万(5472×3648).杉木拍摄时间为 2020 年 11 月 19 日,图像采集天气状况好.拍摄地点位于南平市邵武市下沙镇,地理位置位于 117.60°E、27.33°N,森林覆盖率约 73%,拍摄范围地势较为平坦,相对高差约 100 m,拍摄高度约为 150 m,拍摄对象为杉木人工纯林,拍摄方式为垂直俯视,经过图像预处理、拼接后图像(.tif 格式)如图 1 所示.马尾松拍摄时间为 2020 年 12 月 29 日至 30 日,天气状况良好.拍摄地点位于三明市清流县温郊乡温家山村,拍摄高度约为 180 m,拍摄对象为马尾松人工纯林,拍摄方式为垂直俯视,经过图像预处理、拼接后图像如图 2 所示.

### 1.2 数据标注

采用图像标注工具 LabelImg (<https://github.com/tzutalin/labelImg>),分别从图 1 和图 2 中,找出单一杉木或马尾松区域,采用矩形拉框,长宽大小不一,大约为 224×224,坐标为  $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$ ,杉木标签设为 0,马尾松标签设为 1,共标注了 1 058 张杉木、1 278 张马尾松图像,数据标注完成后生成.xml 文件,该文件只存储矩形拉框的坐标信息,无法直接导出已标注好的图像.

### 1.3 数据裁剪

采用 Python 编程实现批量数据的裁剪.编写

readXML() 函数, 读取 xml 文件坐标, 并转化为 pillow 模块的图像坐标格式 (left, upper, right, lower); 编写 save\_tif() 函数, 从原始的.tif 图像中一次性裁剪出 1058 张杉木、1278 张马尾松图片。



图1 杉木人工纯林



图2 马尾松人工纯林

#### 1.4 数据增强

数据增强是一种能够有效扩充数据集的数据处理手段. 采用 10 种方式对原始数据进行增强: (1) 对原始图像进行垂直翻转和水平翻转; (2) 对原始图像进行逆时针 90°、270°旋转处理; (3) 对原始图像进行灰度化处理; (4) 对原始图像进行 2 种亮化处理 (亮化系数分别为 1.2 和 1.4); (5) 对原始图像进行 2 种饱和度处理 (饱和度系数为 2 和 4); (6) 对原始图像进行随机裁剪, 裁剪后图像大小调整为原始图像大小. 单张马尾松图片数据增强如图 3 所示.

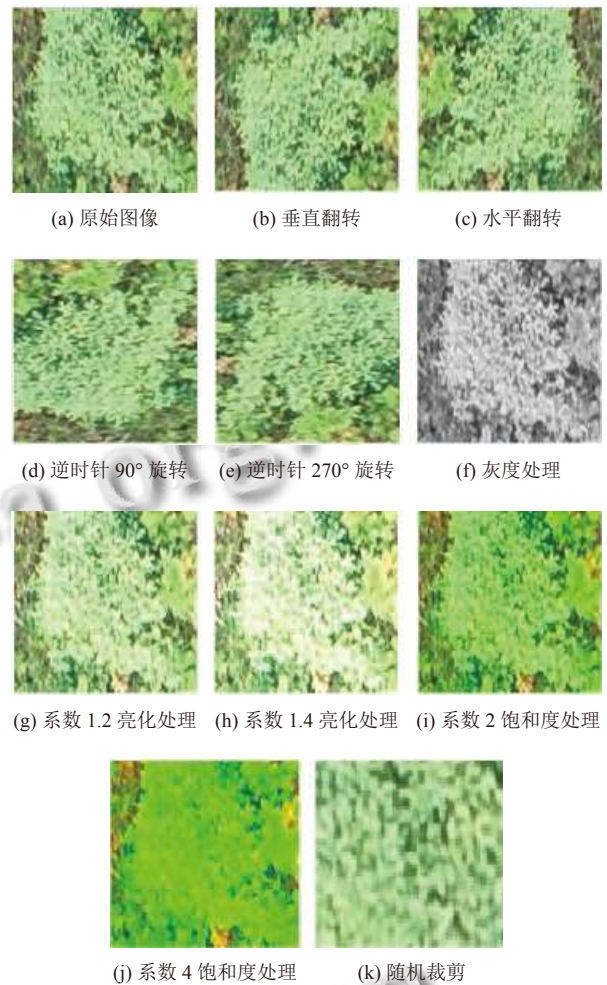


图3 单张马尾松的数据增强

#### 1.5 数据集构建

从数据裁剪后的 1058 张杉木、1278 张马尾松图片中, 选取杉木、张马尾松各 1000 张图片构建数据集, 取名为 UAVTree2k, 其中 2 表示 2 个树种, 2k 表示训练集与测试集的总数为 2000 张, 其中, 杉木和马尾松图片各 1000 张. 定义数据读取函数 read\_image\_filenames(), 读取指定目录下的图片信息, 返回文件名列表和标签列表; 定义解码图片和调整图片大小函数 decode\_image\_and\_resize(), 读取图片并解码, 重新调整图片大小并进行归一化处理; 通过 TensorFlow 框架下的 tf.data.Dataset 的 from\_tensor\_slices() 方法生成小样本数据对 (图片, 标签). 通过数据增强, 将数据裁剪后图像, 扩大为 10580 张杉木图像和 12780 张马尾松图像, 从中选取各 10000 张的图像作为数据集, 取名为 UAVTree20k, 20k 表示训练集与测试集的总数为 20000 张.

## 2 研究方法

### 2.1 VGG16 模型

VGG-Net 是牛津大学计算机视觉组 (visual geometry group) 和 Google DeepMind 公司的研究员一起研发的深度卷积神经网络<sup>[1]</sup>, VGG-Net 在 2014 年大规模视觉识别挑战赛 (ILSVRC) 中, 夺得图像分类第 2 的成绩. VGG-Net 的网络层数从 11 层到 19 层不等. 常用的 VGG16 模型由 13 个卷积层和 3 个全连接层堆叠而

成, 总共包含约 1.38 亿个参数. 所有卷积层使用  $3 \times 3$  的卷积核和 ReLU 激活函数, 以加强特征学习的能力. VGG16 模型整体上分为 8 段, 前 5 段为卷积层, 后 3 段为全连接层. 卷积层重叠 2 或 3 次后形成一个卷积段, 同一段内的卷积层具有相同的卷积核数, 每个卷积段后有一个  $2 \times 2$  的最大池化层, 池化层将图片大小缩小一半. 每增加一个卷积段, 卷积核数增加一倍. VGG16 具体结构见图 4.

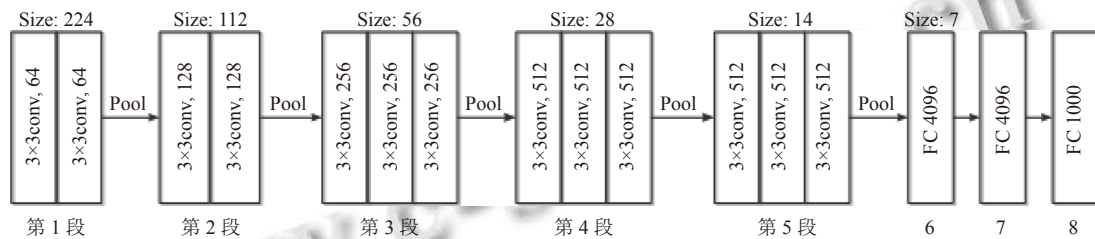


图 4 VGG16 的结构

### 2.2 基于 VGG16 的预训练—微调模型

基于 VGG16 的预训练—微调模型如图 5 所示, 分为共享层和迁移层<sup>[6]</sup>. 共享层为 VGG16 中前 13 个卷积层, 主要提取图像的低级特征. 因为这些特征是通用特征, 所以共享层采用 VGG16 模型在 ImageNet 上训练好的模型, 并冻结其权重与参数. 迁移层提取图像的

高级特征, 是整体模型的核心层. 重新训练 3 个全连接层, 其中包括输出层 FC2, 该层采用 Sigmoid 激活函数. 考虑训练样本较少, 全连接层分别采用 128 和 32 个神经元, 在 FC128 和 FC32 后分别加了 Dropout 方法, 其系数均设置为 0.3, 随机删除全连接层的部分神经元, 防止过拟合, 提高模型泛化能力.

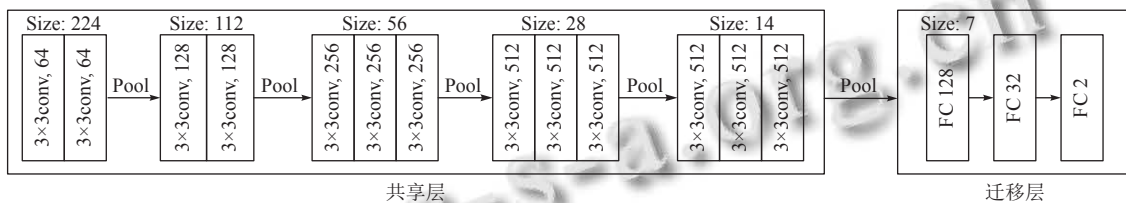


图 5 基于 VGG16 的预训练—微调模型

## 3 实验与结果分析

### 3.1 云环境及其搭建

本机为 Windows 10 的 64 位操作系统, 采用阿里云服务器 2 核 4G 的共享型实例, 在云端搭建 TensorFlow 2.3.0、Anaconda 3.0、Python 3.7 等环境. 采用基于 VGG16 的预训练—微调模型进行无人机可见光影像的树种识别, 数据集采用 UAVTree2k. 基于数据集 UAVTree2k, 进行 VGG16 模型重新训练与测试. 探讨迭代次数、批次大小、数据集划分比例对训练精度与测试精度影响. 所有实验参数更新采用 Adam 优化器, 学习率采用 0.001.

### 3.2 实验结果与分析

#### 3.2.1 迭代次数对精度影响

为探讨迭代次数对精度的影响, 固定训练集 1200 张, 测试集 800 张; 批次大小为 16. 迭代次数依次从 {10, 20, 30, 40, 50, 60} 中取值, 得到训练集的精度分别为 0.9087, 0.9319, 0.9800, 0.9800, 0.9856, 0.9925, 测试集的精度分别为 0.9425, 0.9525, 0.9675, 0.98, 0.9675, 0.9775. 精度与迭代次数关系如图 6 所示. 从图 6 可看出, 训练精度均在 90% 以上, 随着迭代次数的增加, 训练精度整体呈上升趋势; 训练精度均在 94% 以上, 测试精度整体上随迭代次数的增加而增加. 究其原因, 随

迭代次数增加,模型权重、参数不断更新,学习到的特征越多,测度精度也就越好.当迭代次数为40时,测试集精度下降,而后上升,其原因尚不明确.综合训练精度、测试精度,以及训练时间,后续实验的迭代次数取40.

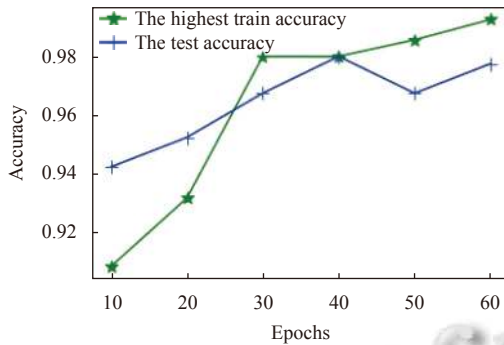


图6 精度与迭代次数关系图

### 3.2.2 批次大小对精度影响

为探讨批次大小对精度的影响,固定训练集1200张,测试集800张;迭代次数为40,批次大小依次从{10, 12, 14, 16}中取值.精度与迭代次数关系如图7所示.从图7可看出,训练精度均在98%以上,随着批次增加,训练精度有些波动.批次大小12时突然下降,批次大小为14时降到最低,此后继续增加;测试精度随批次大小增加而增加,在训练精度下降时,测试精度增速变缓,可能因为批次偏小,训练时反向传播的损失值不稳定导致训练精度的波动,也导致测试精度增速变缓.综合训练精度、测试精度,最优批次大小取16.

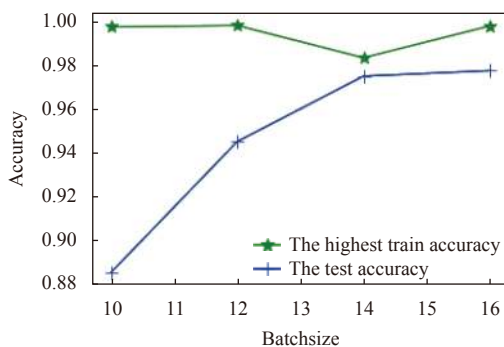


图7 精度与批次大小的关系图

### 3.2.3 不同数据集划分比例对精度影响

为了探讨不同数据集划分比例对精度的影响,验证预训练-微调模型在小样本下的识别效果.训练集与测试集按8:2、7:3、6:4和5:5四种比例划分.这4种

比例的测试精度分别为98%、98%、98.63%、98.20%.4种比例划分训练集的占比依次减少,但实验结果表明准确率有上升趋势,其中6:4的测试精度最高,达到98.63%.原因可能是VGG16模型是在ImageNet上训练过的,共享层的权重和参数进行了冻结,可以较好提取物体识别的通用特征,同时,进一步验证,预训练-微调模型在小样本下,可以取得较好的分类效果.体现了迁移学习具有良好的特征学习能力.

## 4 结论与展望

### 4.1 结论

(1) 论文利用预训练-微调模型,进行无人机可见光影响的树种识别,通过实验研究,在小样本下,模型可以取得较好的分类效果.体现了迁移学习具有良好的特征学习能力.

(2) 模型超参数直接影响模型识别效果,当迭代次数为40,批次大小为16,训练数据与测试数据划分比例为6:4,模型识别效果最好,测试精度达到98.63%.

(3) 利用大疆精灵4RTK无人机,搭载FC6310R相机,采集南平市和三明市的杉木和马尾松人工纯林彩色图像.通过图像预处理、标注、裁剪和增强等环节自主构建两个数据集UAVTree2k和UAVTree20k,分别可用于预训练模型和其他卷积神经网络模型的实验研究.

### 4.2 研究不足与展望

(1) 由于标注工作量大,自主构建的数据集只有2个树种,可进一步采集南方主要树种,把数据集扩充到10树种,每个树种有10000条数据集,最终形成无人机可见光影像的数据集UAVTree,并开放共享,供不同模型与不同算法比较.

(2) 预训练模型是基于ImageNet数据集,是普通图像识别的数据库.而遥感图像的树种识别或无人机影像的树种识别,目标是空对地的识别,因此,预训练模型可采用中国科学院遥感与数字地球研究所建立的全国地表类型遥感影像样本数据集<sup>[12]</sup>,可提高源域和目标域间数据的相似度,通过预训练,把参数调整到合适的范围,从而提高识别效果或至少减少训练模型时间.

(3) 论文基于VGG16预训练-微调模型,只针对数据集进行实验研究,在完成数据集UAVTree时,可对复杂林分下的树种进行分类与制图.同时,非常迫切建

立不同传感器的无人机树种数据集,生成超大规模预训练树种识别模型。

(4) 论文基于云服务器开展实验处理,下一步可进行多 GPU 资源分布式训练方法,如数据并行、模型并行、流水并行等。如何从框架层次自动解决并行策略选择问题是最近的研究热点。

#### 参考文献

- 1 罗仙仙,许松芽,严洪,等.深度学习用于无人机影像树种识别研究进展.泉州师范学院学报,2021,39(2):65-70.
- 2 樊仲谋,周成军,周新年,等.无人机航测技术在森林资源调查中的应用.森林与环境学报,2018,38(3):297-301.
- 3 Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786): 504-507. [doi: 10.1126/science.1127647]
- 4 武红敢,王成波.刍议服务于深度学习研究的林业遥感数据集建设.世界林业研究,2021,34(3):58-62.
- 5 陈琦,刘儒德.教育心理学.3版.北京:高等教育出版社,2020:201.
- 6 王晋东,陈益强.迁移学习导论.北京:电子工业出版社,2021:4-5,119.
- 7 Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge. International Journal of Computer Vision, 2015, 115(3): 211-252. [doi: 10.1007/s11263-015-0816-y]
- 8 王莉影,高心丹.基于迁移学习的遥感影像树种类型分类.实验室研究与探索,2019,38(9):29-33,78. [doi: 10.3969/j.issn.1006-7167.2019.09.008]
- 9 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.激光与电子学进展,2019,56(7):277-286.
- 10 Natesan S, Armenakis C, Vepakomma U. ResNet-based tree species classification using UAV images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, XLII-2/W13: 475-481. [doi: 10.5194/isprs-archives-XLII-2-W13-475-2019]
- 11 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.
- 12 赵理君,郑柯,史路路,等.全国地表类型遥感影像样本数据集.中国科学数据,2019,4(2):195-206.

(校对责编:牛欣悦)