

基于多源数据的毛竹林信息自动提取

王广科¹, 吴达胜^{1,2,3}, 方陆明^{1,2,3}

(1. 浙江农林大学 数学与计算机科学学院, 浙江 杭州 311300; 2. 浙江省林业智能监测与信息技术研究重点实验室, 浙江 杭州 311300; 3. 林业感知技术与智能装备国家林业和草原局重点实验室, 浙江 杭州 311300)

摘要: 毛竹 *Phyllostachys edulis* 是重要的经济林种, 快速准确地获取毛竹林的面积及郁闭度等信息可对毛竹林的高效经营管理提供巨大帮助。基于人工样地的森林资源调查耗时费力且效率低下, 故利用遥感图像等较低成本的数据源估测大范围毛竹林的面积及其郁闭度等信息具有重要意义。本文以浙江省安吉县为研究区域, 基于 Sentinel-2 中等分辨率遥感图像、数字高程模型 (Digital Elevation Model, DEM) 数据及森林资源二类调查数据, 利用 CatBoost、随机森林 (Random Forest, RF)、支持向量机 (Support Vector Machine, SVM) 三种算法, 通过二分类方法估算毛竹林面积和多分类方法估算毛竹林郁闭度。结果表明: 在估算毛竹林面积时, CatBoost、RF、SVM 的总体分类精度分别为 97.75%、96.54%、95.05%, Kappa 系数分别为 0.96、0.93、0.90; 在估算毛竹林郁闭度时, 以上三种算法的总体分类精度分别为 73.38%、73.49%、69.9%, Kappa 系数分别为 0.52、0.52、0.45。在本文研究数据中, 0.7、0.8 郁闭度样本 (大类别样本) 占样本总数的 89.42%, 剩余类别样本 (小类别样本) 只占样本总数的 10.58%。样本失衡是导致毛竹林郁闭度估算精度严重偏低的重要原因。将样本分成两种类别后 (占比较大的样本和占比较小的样本) 分别进行建模, 模型估测效果有了较大改善, 各模型郁闭度估算精度均在 78% 以上, 其中以 CatBoost 最优, 总体精度达到 83.45%。基于 Sentinel-2 遥感图像、DEM 及森林资源二类调查数据的多源数据估算毛竹林面积和郁闭度, CatBoost 算法具有最佳的性能指标, 其估算结果可为竹林资源的监测提供重要借鉴。

关键词: 毛竹林; 遥感; 机器学习; 信息提取

中图分类号: S711 文献标识码: A 文章编号: 1001-3776(2023)03-0079-08

Information Extraction of *Phyllostachys edulis* Forest based on Multi Source Data

WANG Guang-ke¹, WU Da-sheng^{1,2,3}, FANG Lu-ming^{1,2,3}

(1. School of Mathematics and Computer Science, Zhejiang A & F University, Hangzhou 311300, China; 2. Key Laboratory of Forestry Intelligent Monitoring and Information Technology of Zhejiang Province, Hangzhou 311300, China; 3. Key Laboratory of State Forestry and Grassland Administration on Forestry Sensing Technology and Intelligent Equipment, Hangzhou 311300, China)

Abstract: Based on images from Sentinel-2 medium resolution remote sensing imagery, digital elevation model (DEM) and bamboo (*Phyllostachys edulis*) data from forest management survey of Anji county, Zhejiang province in 2018, estimation was made on area and crown density of bamboo forest by CatBoost, random forest (RF) and support vector machine (SVM) model, dichotomy idea and multi-classification. The result demonstrated that classification accuracy of bamboo forest area was 97.75% by CatBoost, 96.54% by RF and 95.05% by SVM, the kappa coefficients were 0.96, 0.93 and 0.9 respectively. The result of Moso bamboo forest canopy density estimated were as follows: the overall classification accuracy of the above three algorithms were 73.38%, 73.49% and 69.9% respectively, and the kappa coefficients were 0.52, 0.52 and 0.45 respectively. The

收稿日期: 2022-12-30; 修回日期: 2023-03-05

基金项目: 浙江省科技重点研发计划资助项目 (2018C02013)

作者简介: 王广科, 硕士, 从事农业工程与信息技术研究; E-mail: 876294612@qq.com。

estimation accuracy of models was greatly improved after the samples divided into two categories (large sample and small sample) and modeled separately. The estimation accuracy of crown density of each model was greater than 78%, especially by CatBoost, 83.45%.

Key words: *Phyllostachys edulis* forest; machine learning; remote sensing; information extraction

作为一种绿色、可再生的森林资源，竹资源对社会经济的可持续发展具有重要作用^[1]。如何快速准确地获取竹林的面积及郁闭度等信息，是对竹资源进行高效经营管理的重要前提。因传统森林资源调查费钱费力且效率低下^[2]，利用遥感图像等较低成本的数据源估测大范围竹资源的面积及其郁闭度等信息成为了近年来的研究热点。

基于遥感图像的竹林分类，经历了监督分类、非监督分类和多源信息复合分类的发展过程^[3]。目前，运用较多的监督学习方法包括决策树(Decision Tree, DT)^[4-7]、随机森林(Random Forest, RF)^[8-9]、支持向量机(Support Vector Machine, SVM)^[10-11]和人工神经网络(Artificial Neural Network, ANN)^[12]等。Du等基于决策树对全球毛竹 *Phyllostachys edulis* 林面积进行估计，总体分类精度达 78.81%^[13]；DIDA等基于随机森林对菲律宾全国竹资源土地覆盖分类绘图，并估计菲律宾共有 14 795 (± 1 283) hm² 竹林^[14]；Wang等使用支持向量机检测中国天目山最近 30 年竹林的扩张情况，得出了该时期竹林总扩张达到 161%，而天目山的总碳储量下降约 4.7%的结论^[15]；WATANABE等利用卷积神经网络对日本竹林进行识别，模型的分类精度在 90%以上，高于传统机器学习模型^[16]。张瑞英等^[17]利用激光雷达等数据建立郁闭度回归模型，在估算温带森林郁闭度时得到了 79.88%的估计精度；邱赛等^[18]利用地球科学激光测高系统(Geoscience Laser Altimeter System, GLAS)波形数据，针对不同森林类型建立森林郁闭度单变量模型和多变量模型，结果显示多变量模型明显优于单变量模型。

目前，森林郁闭度研究多基于回归模型进行反演，但在森林资源调查中，郁闭度被分为了 0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9 九个等级，而这些等级又可作为离散型数据，因此本文运用分类的方法估算毛竹林郁闭度。本文利用 CatBoost、RF、SVM 三种模型，先将研究区的森林林分按是否为毛竹林进行二分类，以此估算毛竹林面积；继而利用上述三种方法，再将估算得到的毛竹林根据郁闭度进行多分类，以此估算毛竹林郁闭度；最后将估算的毛竹林面积及郁闭度与森林资源二类调查的数据进行对比，通过分析以上三种方法的估算精度，确定以上三种方法中最适合提取毛竹林面积及郁闭度信息的模型，将得出的最优估算模型用于提取毛竹林面积及郁闭度，其估算结果可为竹林资源的监测提供重要借鉴。

1 研究数据与方法

1.1 研究区域与研究数据

1.1.1 研究区域 本文选取浙江省安吉县为研究区域(见图 1)。安吉县位于长三角腹地，地理位置为 119°14' ~ 119°53' E、30°23' ~ 30°53' N，年平均气温为 15.43℃，年降水量为 1 496 mm，属于亚热带季风气候。安吉县森林覆盖率达 69.4%，其中竹林面积为 6.7 万 hm²，约占有林地面积的 38%^[19]，丰富的竹林资源为研究提供了良好的条件。

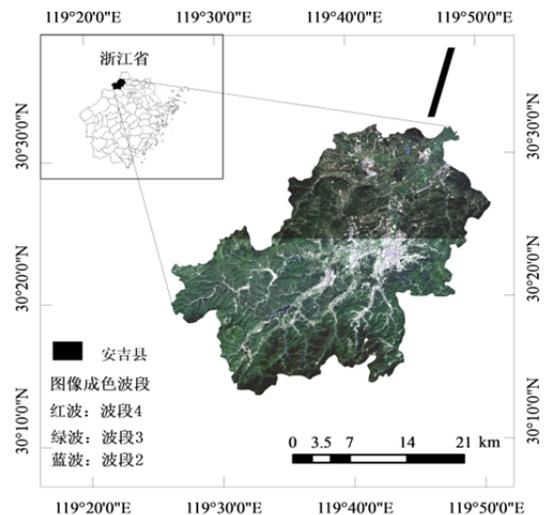


图 1 研究区域 Sentinel-2 遥感影像

Fig. 1 Sentinel-2 remote sensing image of the study area

1.1.2 研究数据 Sentinel-2 携带一枚多光谱成像仪, 可覆盖 13 个光谱波段, 幅宽达 290 km, 地面分辨率为 10 m、20 m 和 60 m。本文从哥白尼开放访问中心获取两幅 Sentinel-2 遥感图像, 图像信息如表 1 所示。DEM 数据是获取高程、坡度、坡向等数据的主要来源, 本文从地理空间数据云获取安吉县 2018 年 10 月的 DEM 数据, 分辨率为 30 m。安吉县毛竹林数据从 2018 年森林资源二类调查中获得, 其中毛竹林与非毛竹林数据统计信息如表 2 所示。

表 1 Sentinel-2 遥感影像基本信息
Tab. 1 Basic information of Sentinel-2 remote sensing image

遥感图像	传感器	等级	感应时间	拼接域编号
Image 1	S2A-MSI	L1C	2018年10月01日	T50RQV
Image 2	S2A-MSI	L1C	2018年10月01日	T50RQU

表 2 安吉县毛竹林与非毛竹林统计数据
Tab. 2 Bamboo forest and non-bamboo forest in Anji county

林分	各等级郁闭度小班数/个							面积/hm ²	小班总数/个
	0.3	0.4	0.5	0.6	0.7	0.8	0.9		
毛竹林	13	18	101	845	4 495	6 634	340	57 614.87	12 446
非毛竹林	33	199	587	2 203	6 150	6 499	1 649	54 518.60	17 320

1.2 研究方法

1.2.1 技术路线 本文的技术路线如图 2 所示。首先, 获取森林资源二类调查、遥感图像及 DEM 数据, 预处理后提取特征因子, 其中, 优势树种、郁闭度为分类标签, 小班面积为辅助计算数据, 其余特征因子用于模型训练; 然后, 利用 CatBoost、RF、SVM 三种模型, 根据优势树种标签将数据分为毛竹林与非毛竹林, 并计算毛竹林面积; 继而利用以上分类结果, 根据郁闭度标签进行多分类。将以上分类结果与森林资源二类调查数据作对比, 计算并分析以上三种模型的分精度。

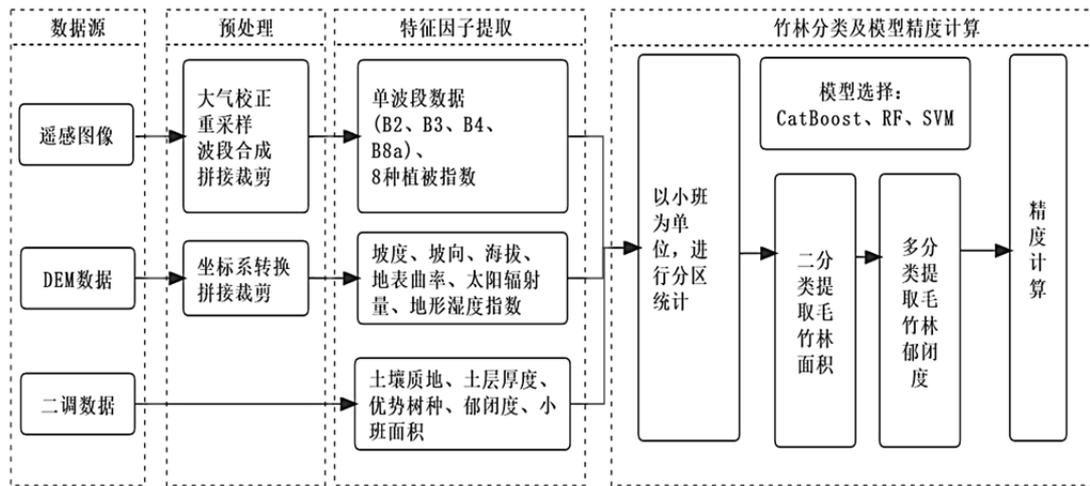


图 2 技术路线

Fig. 2 Technical route

1.2.2 数据预处理 对于 Sentinel-2 遥感图像, 用 Sen2cor (v2.8.0, European Space Agency, French) 进行大气校正后, 通过 SNAP (v8.0, European Space Agency, French) 软件, 采用最近邻插值法 (Nearest Neighbor) 将 30 和 60 m 分辨率的波段分别重采样至 10 m 分辨率。将所有波段重采样至相同分辨率, 不仅能提高图像质量, 而且有利于后续计算提取植被指数。因为森林资源二类调查数据为西安 80 坐标系, 所以用 ENVI (v5.3, Exelis Visual Information Solutions, US) 软件将 Sentinel-2 遥感图像和 DEM 图像由 WGS_1984 坐标系转为西安 80 坐标系, 然

后对遥感图像根据安吉县行政边界进行拼接裁剪处理。

1.2.3 特征因子选择 提取遥感图像 B2、B3、B4、B8a 单波段数据，并通过以上波段数据计算植被指数。本文提取 8 种植被指数^[20]，其计算公式如表 3 所示。

表 3 8 种植被指数的计算公式
Tab. 3 Formulae of the eight vegetation indices

指数名称	计算公式	指数名称	计算公式
归一化植被指数 (NDVI)	$NDVI = (NIR - R) / (NIR + R)$	归一化水指数 (NDWI)	$NDWI = (G - NIR) / (G + NIR)$
土壤调整植被指数 (SAVI)	$SAVI = 1.5 \times (NIR - R) / 0.5 \times (NIR + R)$	类胡萝卜素指数 (CRI)	$CRI = 1/B - 1/G$
差值植被指数 (DVI)	$DVI = NIR - R$	非线性植被指数 (NVI)	$NVI = (NIR^2 - R) / (NIR^2 + R)$
增强型植被指数 (EVI)	$EVI = 2.5(NIR - R) / (NIR + 6.0 \times R - 7.5 \times B + 1)$	归一化差异绿度指数 (NDGI)	$NDGI = (G - R) / (G + R)$

注: NIR 表示近红外波段 B8a, R 表示红波段 B2, G 表示绿波段 B3, B 表示蓝波段 B4。

从 DEM 数据提取海拔、坡度、坡向、地表曲率、太阳总辐射量、地形湿度指数 (Topographic Wetness Index, TWI) 6 个特征因子。地表曲率可基于四次表面模型在 ArcGIS (v10.4.1, Environmental Systems Research Institute, US) 软件的 spatial analysis surface 模块中直接提取^[21]。通过 ArcGIS Spatial Analyst 模块中的太阳辐射分析工具计算太阳总辐射量, 其计算根据 Rich 等^[22]创立并进一步由 Fu 等^[23]发展的半球视域算法, 即太阳总辐射量等于所有太阳图和星空图扇区的直接辐射量和散射辐射量相加。目前应用最广泛的地形湿度指数 (TWI) 计算方法^[24], 其计算公式为: $TWI = \ln(x/\tan y)$, 其中 x 为单位等高线长度上的汇水面积, y 为局部坡度。

从安吉县 2018 年森林资源二类调查数据中提取土层厚度、土壤质地、优势树种、郁闭度、面积 5 个特征因子。毛竹林生长状况与土壤条件存在相关性, 土层厚度、土壤质地等立地条件对毛竹林生长有显著性影响^[25]。优势树种、郁闭度为分类标签, 面积为验证数据, 分类后计算毛竹林和非毛竹林面积, 和森林资源二类调查数据对比并以此评判模型性能。

1.2.4 分类与评估方法 将选择的特征因子以森林资源二类调查的小班为单元进行分区统计。分区统计后共 43 212 条数据, 每条数据包含 14 个特征因子、小班面积以及优势树种、郁闭度分类标签。删除因小班为未成林等因素导致平均胸径为 0 的 13 444 条数据, 将剩余的 29 768 条数据按 70%、15% 和 15% 的比例随机划分为训练集、验证集和测试集, 按优势树种是否为毛竹进行二分类 (即划分为毛竹林和非毛竹林两大类)。在上述 29 768 条数据中, 有 12 446 条毛竹林数据, 将毛竹林数据按各等级郁闭度样本数的 70%、15%、15% 比例, 随机划分为训练集、验证集和测试集, 根据郁闭度标签进行多分类。

本文选用 CatBoost、RF、SVM 三种分类模型。CatBoost 是最近流行的一种梯度增强算法, 它以回归型决策树为基学习器, 通过梯度增强 (Gradient Boosting, GB) 来迭代拟合强学习^[26]。在验证集上通过十折交叉验证, 根据验证结果绘制接收者操作特征曲线 (Receiver Operating Characteristic Curve, ROC), 并计算接收者操作特征曲线下的面积 (Area Under Curve, AUC)。由实验结果发现, 根据 CatBoost 默认参数训练的模型就能取得很好的 AUC 评分, 调参后模型 AUC 评分没有明显提高。

RF 是典型的集成学习算法, 它通过对训练样本有放回的抽样形成训练子集, 再从训练子集中抽取特征因子进行基学习器的构建^[27]。用十折交叉法进行验证, 其中树木数量范围从 50 棵到 500 棵, 步长为 5; 最大树木深度范围从 10 层到 200 层, 步长为 5。当树木数量为 135 棵、最大树深为 35 层时, 模型 AUC 评分最高。SVM 是把训练集非线性映射到高维特征空间, 通过高维特征空间建立最优分离超平面, 从而将训练集根据最优分离超平面划分为不同的集合^[28]。用十折交叉法进行验证, 正则化参数范围从 1 到 10, 步长为 1; 内核系数分别为 0.01、0.03、0.05、0.1、0.3、0.5、1、3、5。本文采用径向基核函数 (Radial Basis Function, RBF), 当正则化参数为 3.0、内核系数为 0.1 时, 模型的交叉熵损失函数获得最小值, AUC 评分最高。为比较不同算法的分类准确性, 本文选用的模型性能指标有总体分类精度 (Overall Accuracy, OA)、用户精度 (User Accuracy, UA)、生产者精度 (Producer Accuracy, PA) 以及 Kappa 系数。

2 结果与分析

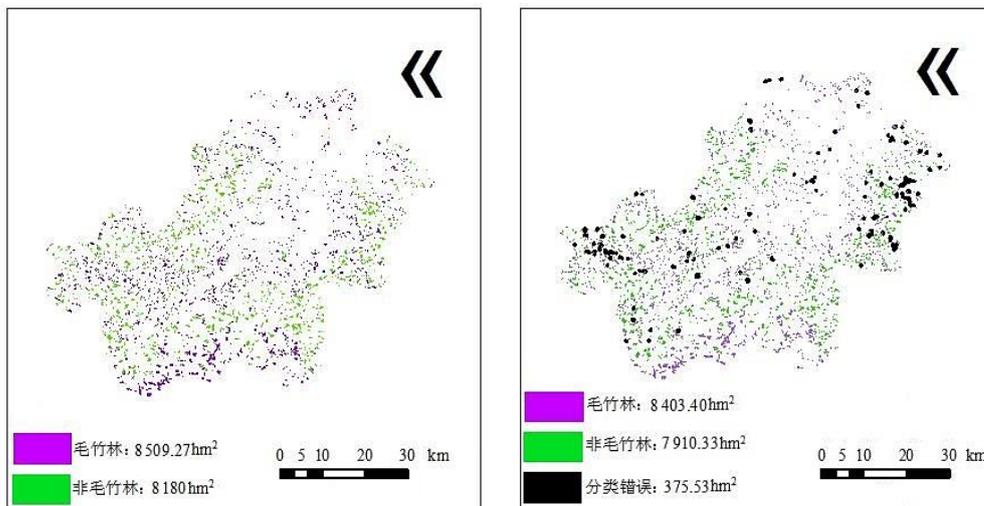
2.1 毛竹林提取与面积估算

先训练 CatBoost、RF、SVM 二分类模型, 然后把测试集按优势树种分为毛竹林和非毛竹林, 继而把分类结果与森林资源二类调查结果进行对比, 计算以上三种模型的性能指标, 结果见表 4。

表 4 二分类模型性能指标
Tab. 4 Performance indicators for binary classification models

类别	CatBoost		RF		SVM	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
毛竹林	96.89	98.76	94.88	98.54	93.74	96.74
非毛竹林	98.68	96.70	98.42	94.47	96.50	93.28
OA/%	97.75		96.54		95.05	
Kappa	0.96		0.93		0.90	

由表 4 可知, CatBoost 模型的总体分类精度最高, 为 97.75%。其中, 测试集毛竹林的实测面积为 8 509.27 hm², 非毛竹林的实测面积为 8 180 hm² (如图 3A 所示), CatBoost 估算的毛竹林面积为 8 403.40 hm²、非毛竹林面积共 7 910.33 hm² (如图 3B 所示), 错分面积共 375.53 hm²。



A - 测试集中毛竹林和非毛竹林的实际分布

B - 基于 CatBoost 的毛竹林分布估算

图 3 毛竹林分布

Fig. 3 Distribution of bamboo forest

2.2 毛竹林郁闭度估算

先训练 CatBoost、RF、SVM 多分类模型, 然后把测试集数据按毛竹林郁闭度等级进行多分类, 进而对比模型分类的郁闭度等级与森林资源二类调查的郁闭度等级, 计算以上三种模型的性能指标 (表 5)。由表 5 可知, 在估算毛竹林郁闭度时, 对于郁闭度 0.6、0.7 两个类别的数据, CatBoost 模型的 UA、PA 评分均比 RF 评分高; 而对于郁闭度 0.8、0.9 两个类别的数据, CatBoost 模型的 UA、PA 评分均低于 RF 评分。但总体来看, CatBoost 和 RF 模型的 Kappa 系数相同, RF 模型的总体分类精度最高, 为 73.49%。测试集中毛竹林郁闭度的实际分布 (来自森林资源二类调查数据) 如图 4A 所示, 基于 RF 的郁闭度估算结果如图 4B 所示, 两者具有较好的一致性,

体现出了基本相同的分布趋势：安吉县毛竹林郁闭度南高北低、四周高中间低。

表 5 多分类模型性能指标
Tab. 5 Performance indicators for multi-classification model

郁闭度	CatBoost		RF		SVM	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
0.3	0	0	0	0	0	0
0.4	0	0	0	0	0	0
0.5	0	0	0	0	0	0
0.6	67.53	40.94	67.14	37.01	59.04	38.58
0.7	69.65	68.10	69.20	67.66	65.80	63.65
0.8	76.08	85.33	76.35	86.63	73.12	82.31
0.9	71.43	19.61	100.00	13.73	72.73	15.69
OA/%	73.38		73.49		69.90	
Kappa	0.52		0.52		0.45	

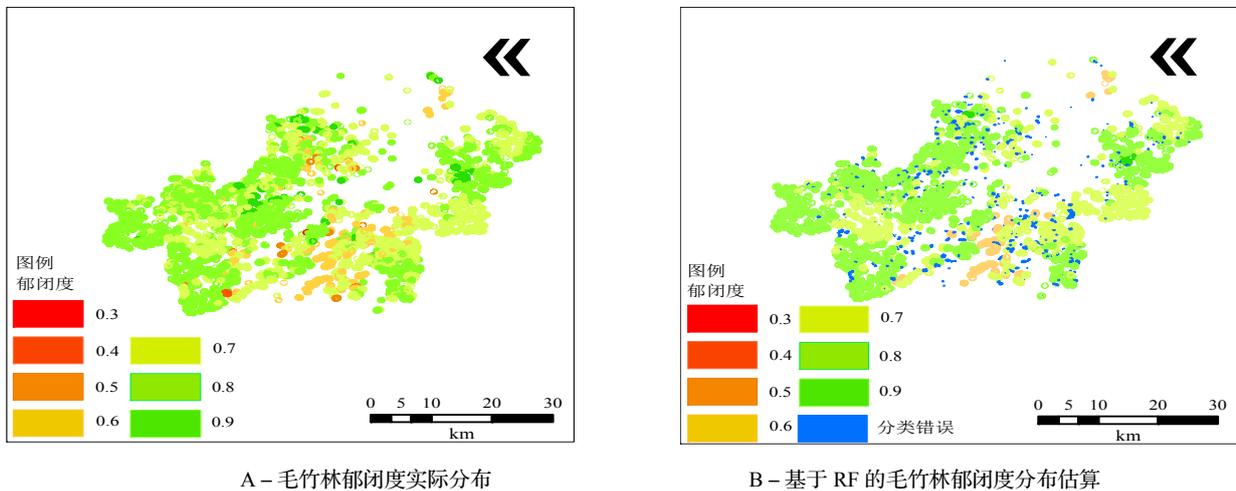


图 4 毛竹林郁闭度分布

Fig. 4 Crown density distribution of bamboo forest

2.3 毛竹林郁闭度估算结果分析

由表 4 可知，在二分类模型中，CatBoost、RF、SVM 的总体分类精度、Kappa 系数均依次减小，且三种模型的总体分类精度均在 95.00% 以上，Kappa 系数也均在 0.90 以上，分类效果较好。但由表 5 可知，多分类模型的总体分类精度远低于二分类模型（分别低 24.37%、23.05% 和 25.15%），即使是总体分类精度最高的 RF 模型也仅为 73.49%。

2.4 多分类模型估算郁闭度

基于 CatBoost、RF、SVM 三种模型对毛竹林郁闭度的估算精度显著低于毛竹林面积的估算精度，样本失衡可能是导致该结果的一个重要原因。在二分类中，毛竹林与非毛竹林数据分别为 12 446 和 17 322 条，分别占训练集样本总量的 41.81%、58.19%，分布相对平衡。在多分类中，毛竹林郁闭度有 0.3、0.4、0.5、0.6、0.7、0.8、0.9 七类，各类样本分别占训练集总样本量的 0.10%、0.14%、0.81%、6.79%、36.12%、53.30%、2.73%。其中郁闭度为 0.8 的样本量最大，占训练集总样本量的 53.30%，郁闭度为 0.3 的样本量最小，仅占训练集总样本量的 0.10%，样本类别分布极不均匀，在训练模型时容易导致欠拟合，造成模型总体分类精度偏低。

本文尝试改进数据集和分样本建模。在改进数据集方面，尝试采用 Scikit-learn 官方提供的 balanced 类别加权方法对数据集进行平衡处理，但模型总体分类精度不增反降（如表 6 所示），分析其根源可能在于：虽然类别加权增强了少量样本，但同时也增大了大样本的损失函数，故最终导致模型总体分类精度下降^[29]。分样本建

模即根据样本量占比大小分类建模, 本文将占比较大的郁闭度 0.7 和 0.8 两类样本合并建模(大类别样本模型)、剩余占比较小的样本类别合并进行建模(小类别样本模型), 再将上述两种模型分别对测试集进行估测, 结果表明, 分类建模可显著提高模型分类精度(如表 6 所示), 其中 CatBoost、RF 模型最终总体分类精度均在 80% 以上, SVM 模型的总体分类精度也达到了 78.2%。

表 6 基于多分类模型的毛竹林郁闭度估测性能指标
Tab. 6 Performance indicators of multi classification model for estimating crown density of bamboo forest

模型	CatBoost		RF		SVM		
	OA/%	Kappa	OA/%	Kappa	OA/%	Kappa	
原始模型	73.38	0.52	73.49	0.52	69.90	0.45	
类别加权	66.74	0.47	71.99	0.48	61.65	0.40	
分样本建模	大类别样本模型	83.58	0.66	82.14	0.63	78.13	0.54
	小类别样本模型	82.32	0.61	83.84	0.64	78.79	0.53
	最终总体性能	83.45	0.71	83.18	0.71	78.20	0.62

3 结论与讨论

本文基于 Sentinel-2 遥感图像数据、DEM 数据、森林资源二类调查数据, 利用 CatBoost、RF、SVM 三种算法估算毛竹林面积及其郁闭度。结果表明, 在毛竹林面积估算中, CatBoost 模型的总体分类精度最高, 为 97.75%; 在估算毛竹林郁闭度时, RF 模型的总体精度最高, 为 73.49%。

当估测毛竹林郁闭度时, 基于总体样本进行统一估测时, 因样本极度失衡导致估测精度不高(最高精度仅为 73.49%)。当采用 Scikit-learn 官方提供的 balanced 类别加权方法对数据集进行平衡处理时, 模型总体分类精度不增反降。而采用分类建模方法, 将样本分成两种类别(占比较大的样本和占比较小的样本)时, 模型估测精度有了显著提升, 各模型郁闭度估算精度均在 78% 以上, 以 CatBoost 最优, 总体分类精度达到 83.45%。从算法而言, CatBoost 更擅长于自动降低噪声数据的影响, 小样本数据在训练集中相当于噪声点, 而 CatBoost 采用排序提升法对抗训练集中的噪声点, 从而减少梯度估计的偏差, 进而解决预测偏移问题, 因此在上述三种模型中 CatBoost 的总体分类精度最高, 可为毛竹林的低成本监测提供借鉴。

本研究利用单一时相遥感图像进行分类, 无法充分保证毛竹信息的时效性、动态性, 今后的研究可利用大尺度、多时相遥感图像动态监测毛竹林变化信息。此外, 针对多分类样本失衡问题, 本文运用的分类建模方法, 是通过数据层面对模型进行改进, 是否能找到处理样本失衡的更好方法, 仍有待进一步研究。

参考文献:

- [1] MANANDHAR R, KIM J H, KIM J T. Environmental, social and economic sustainability of bamboo and bamboo-based construction materials in buildings[J]. *J Asian Arch Build Engin*, 2019, 40 (24): 9541 - 9555.
- [2] 魏芬娟, 陈崇成, 叶晓燕, 等. 森林资源调查与督查的移动 GIS 研发及应用[J]. *计算机工程与应用*, 2020, 56 (6): 231 - 238.
- [3] 蒲梦馨, 倪忠云, 赵银兵, 等. 竹林信息遥感提取方法研究进展[J]. *四川林业科技*, 2021, 42 (4): 122 - 131.
- [4] 高国龙, 杜华强, 韩凝, 等. 基于特征优选的面向对象毛竹林分布信息提取[J]. *林业科学*, 2016, 52 (9): 77 - 85.
- [5] 吴梓尚, 林辉, 孙华, 等. 基于 MODIS 影像的森林类型决策树分类方法研究[J]. *中南林业科技大学学报*, 2017, 37 (9): 80 - 85.
- [6] 杜华强, 孙晓艳, 韩凝, 等. 综合面向对象与决策树的毛竹林调查因子及碳储量遥感估算[J]. *应用生态学报*, 2017, 28 (10): 3163 - 3173.
- [7] 崔璐, 杜华强, 周国模, 等. 决策树结合混合像元分解的中国竹林遥感信息提取[J]. *遥感学报*, 2019, 23 (1): 166 - 176.
- [8] 熊恬苇. 江西省竹林时空变化遥感研究[D]. 南昌: 江西师范大学, 2018.
- [9] 严欣荣, 张美曼, 郑亚雄, 等. 基于 Sentinel-2 的丛生竹林信息提取方法比较及分布特征[J]. *生态学杂志*, 2020, 39 (3): 1056 - 1066.
- [10] 徐辉, 潘萍, 杨武, 等. 基于多源遥感影像的森林资源分类及精度评价[J]. *江西农业大学学报*, 2019, 41 (4): 751 - 760.
- [11] 钟滨, 廖永皓, 蔡海生. 基于光谱与纹理特征的高分二号竹林信息提取以庐山自然保护区为例[J]. *测绘与空间地理信息*, 2020, 43 (12): 8 - 13.
- [12] 王聪. 基于多源遥感数据的毛竹林冠层郁闭度多尺度反演研究[D]. 杭州: 浙江农林大学, 2015.

- [13] DU H Q, MAO F J, LI X J, et al. Mapping global bamboo forest distribution using multisource remote sensing data[J]. IEEE J SelcTop App Earth Observ Remot Sens, 2018, 11 (5) : 1 – 14.
- [14] DIDA J J V, Araza A B, Eduarte G T, et al. Towards nationwide mapping of bamboo resources in the Philippines: testing the pixel-based and fractional cover approaches[J]. Int J Remot Sens, 2021, 42 (9) : 3380 – 3404.
- [15] WANG Y, JIN J X, JIANG H, et al. Satellite-based detection of bamboo expansion over the past 30 years in Mount Tianmushan, China[J]. Int J Remot Sens, 2016, 37 (13 – 14) : 2908 – 2922.
- [16] WATANABE S, SUMI K, ISE T. Identifying the vegetation type in Google Earth images using a convolutional neural network: a case study for Japanese bamboo forests[J]. BMC Ecol, 2020, 20 (1) : 65 – 79
- [17] 张瑞英, 庞勇, 李增元, 等. 结合机载 LiDAR 和 LANDSAT ETM+数据的温带森林郁闭度估测[J]. 植物生态学报, 2016, 40 (2) : 102 – 115.
- [18] 邱赛, 邢艳秋, 田静, 等. 基于 ICESat-GLAS 波形数据估测森林郁闭度[J]. 南京林业大学学报 (自然科学版), 2016, 59 (05) : 99 – 106.
- [19] 杜华强, 周国模, 葛宏立, 等. 基于 TM 数据提取竹林遥感信息的方法[J]. 东北林业大学学报, 2008, 36 (3) : 35 – 38.
- [20] 蔡林菲, 吴达胜, 方陆明, 等. 基于 XGBoost 的高分二号影像树种识别[J]. 林业资源管理, 2019, 0 (5) : 44 – 51.
- [21] 李倩楠. 基于不同算法的 DEM 地面曲率提取的比较分析[J]. 首都师范大学学报 (自然科学版), 2016, 37 (5) : 82 – 85.
- [22] RICH P M, DUBAYAH R, HETRICK W A, et al. Using viewshed models to calculate intercepted solar radiation: applications in ecology[C]. Proceedings of American society for photogrammetry and remote sensing technical. Bethesda: American Society for Photogrammetry and Remote Sensing, 1994: 524 – 529.
- [23] FU P, RICH P M. A geometric solar radiation model with applications in landscape ecology[D]. Lawrence Kansas: University of Kansas, 2000.
- [24] QUINN P, BEVEN K J, LAMB R. The $\ln(\alpha/\tan\beta)$ index: how to calculate it and how to use it within the TOPMOD-EL framework[J]. Hydrol Proc, 2010, 9 (2) : 161 – 182.
- [25] 骆爱琴, 曾玲梅, 王勤. 安徽铜陵毛竹林生长与土壤特点分析[J]. 山东林业科技, 2010, 40 (2) : 40 – 41.
- [26] IBRAHIM A A, RIDWAN R L, MUHAMMED M M, et al. Comparison of the catBoost classifier with other machine learning methods[J]. Int J Adv Comput Sci App, 2020, 11 (11) : 11.
- [27] 董师师, 黄哲学. 随机森林理论浅析[J]. 集成技术, 2013, 2 (1) : 1 – 7.
- [28] 张浩然, 韩正之, 李昌刚. 支持向量机[J]. 计算机科学, 2002, 29 (12) : 135 – 137, 142.
- [29] THOMAS J, JOUVE P E, NICOLYANNIS N. Optimisation and evaluation of random forests for imbalanced datasets[J]. Lect Note Artif Intell, 2006, 622 – 631.