

DOI:10.12171/j.1000-1522.20190269

基于 CART 决策树的自然保护区植被类型分布研究

白莹¹ 胡淑萍²

(1. 国家林业和草原局信息中心, 北京 100714; 2. 中国林业科学研究院资源信息研究所, 北京 100091)

摘要:【目的】针对保护区监测需求, 充分发挥 GF-1 WFV 影像的宽幅特点和面向对象、机器学习算法在遥感影像分类中的优势, 提高保护区植被类型遥感监测的精度, 为保护区管理决策提供依据。【方法】以甘肃省白水江国家级自然保护区为研究区, 主要数据源包括 GF-1 WFV 多光谱数据、Landsat-8 OLI 遥感数据、DEM 数据、野外调查数据等。首先, 对 GF-1 WFV 数据进行多尺度分割, 将研究区划分为诸多区域性的分割对象; 然后, 以分割对象为基本单元, 研究光谱特征、几何特征、纹理特征不同组合情况下, 基于 CART 决策树分类的结果; 最后, 利用训练样本建立基于 TTA 的精度检验, 并基于混淆矩阵对分类结果进行分析。【结果】在多尺度分割过程中, 形状因子、紧致度分别设置为 0.2 和 0.5 时地物边界显示较好; 当形状因子和紧致度固定时, 研究区最佳分割尺度为 40。精度检验结果表明, 基于 CART 决策树的保护区植被类型分类结果整体精度均在 83% 以上, Kappa 系数在 0.80 以上, 优于最邻近分类法和支持向量机分类算法, 其中基于光谱特征、几何特征、纹理特征的 CART 决策树分类结果精度最高, 总体精度为 85.18%, Kappa 系数为 0.832 2, 优于光谱特征分类、光谱特征结合几何特征分类的方法。【结论】基于 CART 决策树算法和面向对象方法的 GF-1 遥感影像分类方法适用于保护区植被类型分布研究, 可有效辅助保护区监测工作。

关键词: 面向对象; CART 决策树; 植被类型; GF-1 遥感数据; 自然保护区

中图分类号: S771.8 文献标志码: A 文章编号: 1000-1522(2020)06-0113-10

引文格式: 白莹, 胡淑萍. 基于 CART 决策树的自然保护区植被类型分布研究 [J]. 北京林业大学学报, 2020, 42(6): 113-122. Bai Ying, Hu Shuping. Vegetation type distribution in nature reserve based on CART decision tree [J]. Journal of Beijing Forestry University, 2020, 42(6): 113-122.

Vegetation type distribution in nature reserve based on CART decision tree

Bai Ying¹ Hu Shuping²

(1. Information Center of National Forestry and Grassland Administration, Beijing 100714, China;

2. Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: [Objective] In order to meet the needs of nature reserve monitoring, improve the precision of remote sensing inventory on vegetation types, an object-oriented classification method and machine learning algorithms were developed in vegetation classification by GF-1 WFV remote sensing data. [Method] The study site is located in Baishuijiang National Natural Reserve, Gansu Province of northwestern China. The GF-1 WFV multispectral data, Landsat-8 OLI remote sensing data, DEM data and field survey data were employed as the key data sources. Firstly, the multiresolution segmentation of GF-1 WFV data was carried out, and the research area was divided into many polygon objects. Then spectral features, geometric features and texture features from polygon objects were extracted to vegetation classification using CART decision

收稿日期: 2019-06-26 修回日期: 2019-09-04

基金项目: 国家林业和草原局大熊猫国际合作资金项目(2017115IFRIT01)。

第一作者: 白莹, 工程师。主要研究方向: 林业信息技术应用。Email: baiying1988@126.com 地址: 100714 北京市东城区和平里东街 18 号国家林业和草原局信息中心。

责任作者: 胡淑萍, 博士, 助理研究员。主要研究方向: 林业遥感和地理信息系统应用。Email: hushuping@ifrit.ac.cn 地址: 100091 北京市海淀区东小府 1 号中国林科院资源信息所。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

tree. Finally, the accuracy of classification was analyzed by error matrix based on TTA mask. [Result] In the multiresolution segmentation process, the shape factor and compactness were set to 0.2 and 0.5, respectively, the boundary of the polygon objects was identical with ground objects. When the shape factor and compactness were fixed, the optimal segmentation scale was 40. The accuracy results showed that the overall accuracy and Kappa coefficient were exceed of 83% and 0.80 in three CART decision trees, which was superior to KNN algorithm and SVM algorithm. Overall accuracy and Kappa coefficient of CART decision tree reached 85.18% and 0.832 2 by spectral, geometric and texture features, which was better than CART decision tree by spectral features or spectral combined with geometric features. [Conclusion] The image classification based on CART decision tree algorithm and object-oriented classification method were suitable for vegetation classification in nature reserve by GF-1 WFV image, which could effectively assist the nature reserve monitoring.

Key words: object-oriented; CART decision tree; vegetation type; GF-1 remote sensing data; nature reserve

自然保护区是指对有代表性的自然生态系统、珍稀濒危野生动植物物种的天然集中分布、有特殊意义的自然遗迹等保护对象所在的陆地、陆地水域或海域,依法划出一定面积予以特殊保护和管理的区域。建立自然保护区是保护生物多样性及其生境的重要途径之一^[1-4],也是最有效可行的就地保护手段^[5]。植被是陆地生态系统的主体,在维持生态平衡、生态过程以及生态服务功能等方面发挥着不可替代的作用,所以分析、监测、评价植被组成与变化已成为研究保护区生态环境状况最为有效的手段。

高空间分辨率遥感影像是植被类型识别的最佳数据源之一,其光谱、几何、纹理特征有助于提高森林类型的识别能力^[6],但该类影像获取成本高,单景影像覆盖面积小,卫星回访周期长,难以满足大面积保护区监测工作需求。高分一号(GF-1)遥感卫星是我国民用高分系列卫星的首发星,该卫星配置了两台PMS相机,可获取8 m多光谱和2 m全色影像,覆盖周期为41 d;该卫星还配置了4台多光谱宽幅相机,幅宽可达200 km,空间分辨率为16 m,覆盖周期为4 d,其较高的时间分辨率和宽幅覆盖能力能有效补充研究区多云雨天气下光学影像的不足,对大中尺度保护区监测具有良好的应用潜力。

常见的面向对象分类方法包括两种:一种是基于规则的分类,包括阈值分类、隶属度函数分类;另一种是机器学习分类,包括最邻近分类、贝叶斯分类、支持向量机分类、随机森林分类等。其中决策树分类在影像分类和植被提取中得到广泛应用。Sun等^[7]基于多时相Landsat数据,利用决策树算法提取西双版纳橡胶的分布;郝泷等^[8]运用Landsat-8影像的纹理特征等,基于决策树算法等提取西藏林芝县的森林地物类别;雷光斌等^[9]利用多源多时相遥感影像,结合决策树算法提取岷江上游山地森林

植被类型;钱军朝等^[10]基于WorldView-2影像,采用决策树分类算法提取了杭州西湖区的绿地信息。在面向对象分类方面,不同研究使用不同的决策树实现了影像分类,但针对GF-1号影像的CART决策树植被类型分类研究则鲜见。

本研究以甘肃白水江国家级自然保护区为研究对象,以面向对象的软件eCognition为植被类型提取的平台,比较不同尺度分割的效果,研究光谱特征、几何特征、纹理特征不同组合下的识别精度,并利用CART决策树进行分类研究,掌握研究区植被类型分布,为保护区监测管理提供支撑。

1 研究区概况与数据处理

1.1 研究区概况

甘肃白水江国家级自然保护区位于甘肃省陇南市(图1),地理坐标为:104°16'~105°25' E, 32°36'~33°00' N。地势西北高、东南低,最高海拔4 072 m,最低海拔550 m,总面积1 867.75 km²。河流属嘉陵江水系,是北亚热带气候与暖温带气候的过度地区。相对湿度50%~70%,年平均日照1 700 h,平均气温15℃,降雨多集中在6—9月,约占全年降水量的75%~80%。该区有高等植物197科2 160种,主要优势树种有岷江冷杉(*Abies faxoniana*)、麦吊云杉(*Picea brachytyla*)、红桦(*Betula albosinensis*)、白桦(*B. platyphylla*)、辽东栎(*Quercus wutaishansea*)、栓皮栎(*Q. variabilis*)、米心水青冈(*Fagus engleriana*)、房县枫(*Acer sterculiaceum* subsp. *franchetii*)等。

1.2 研究数据及预处理

1.2.1 遥感影像数据

研究主要以2016年6月16日的GF-1 WFV影像为研究对象,影像来源是资源卫星应用中心(<http://www.cresda.com/>),其多光谱波段空间分辨率为16 m,包括4个波段:蓝光波段(430~550 nm)、

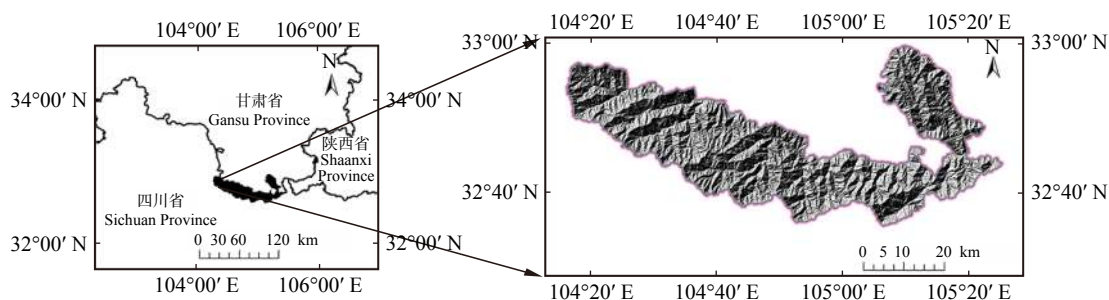


图1 研究区地理位置图

Fig. 1 Location of the study area

绿光波段(490 ~ 610 nm)、红光波段(600 ~ 720 nm)和近红外波段(750 ~ 950 nm)。

Landsat-8 OLI 影像下载自地理空间数据云(<http://www.gscloud.cn/>), 成像时间为2016年7月9日, 影像

包括空间分辨率为30 m的8个波段(433 ~ 2 300 nm)和空间分辨率15 m的全色波段(500 ~ 680 nm), 该影像全色波段用于辅助GF-1 WFV影像进行几何校正。影像基本信息见表1。

表1 研究区遥感影像信息

Tab. 1 Image acquisition information in the study area

影像标识符 Image identifier	传感器类型 Sensor type	轨道号 Track No. (path/row)	太阳高度角 Solar altitude angle/(°)	太阳方位角 Solar azimuth/(°)	云量 Cloud percentage/%
GF1_WFV4_E105.1_N33.5_20160616_L1A0001650652	GF1-WFV4	17/100	78.403	149.749	0
LC81290372016183LGN00_MTL	Landsat-8 OLI	129/37	67.789	111.247	3.68

1.2.2 地面调查数据

为辅助影像分类, 项目组于2018年8月22日至9月14日在研究区开展野外调查工作。利用华测LT500H手持终端采集样点地理坐标, 记录样点地物覆盖类型并拍摄照片, 得到野外调查样本740个。外业调查数据与遥感影像相差2年, 但保护区内干扰较少, 因此认为该数据能反映2016年保护区植被覆盖情况。同时辅以部分第4次大熊猫调查数据, 补充各植被类型样点数据318个。

1.2.3 其他辅助数据

研究区DEM数据下载自地理空间数据云(<http://www.gscloud.cn/>), GDEMDEM空间分辨率为30 m, 用于GF-1 WFV影像的正射校正。Google Earth影像用以辅助样本选择, 尤其是非林地样本数据。森林资源二类调查小班资料用于分类系统制定。

1.2.4 遥感影像预处理

GF-1 WFV影像采用ENVI 5.3软件进行辐射定标、大气校正、正射校正等预处理。辐射定标采用中国资源卫星应用中心提供的绝对辐射定标系数进行校正; 大气校正采用FLAASH大气校正模型; 正射校正采用影像自带的PRC(rational polynomial coefficient)文件和30 m空间分辨率的DEM进行校正; 几何校正利用image registration workflow工具, 以Landsat-8 OLI影像全色波段(15 m空间分辨率)

为基准影像, 正射校正后的GF-1 WFV影像为待校正影像, 重采样方法选用3次卷积法。

2 研究方法

2.1 技术路线

本研究采用面向对象的监督分类方法, 主要步骤包括对GF-1 WFV影像进行预处理、影像分割、确定最佳分割尺度、以分割单元为基本对象建立训练样本、组合光谱特征、几何特征、纹理特征构建CART决策树分类器、与最邻近分类法(KNN)和支持向量机(SVM)分类法进行精度对比等步骤。具体技术流程见图2。

2.2 分类系统的制定

根据《土地利用现状分类》(GB/T21010—2007)国家标准和研究区的实际情况, 将研究区一级土地类型划分为耕地、林地、居民区、道路、水体、裸岩6类。从研究区森林资源二类调查小班资料分析可知, 该区主要植被类型有针叶林、针阔混交林、阔叶林、灌木林和竹林5类。综合实际情况, 本研究制定的分类系统共二级10个类型, 即: 耕地、居民区、道路、水体、裸岩、针叶林、针阔混交林、阔叶林、灌木林和竹林。

2.3 样本获取

分类器训练和精度检验样本基于外业调查数据

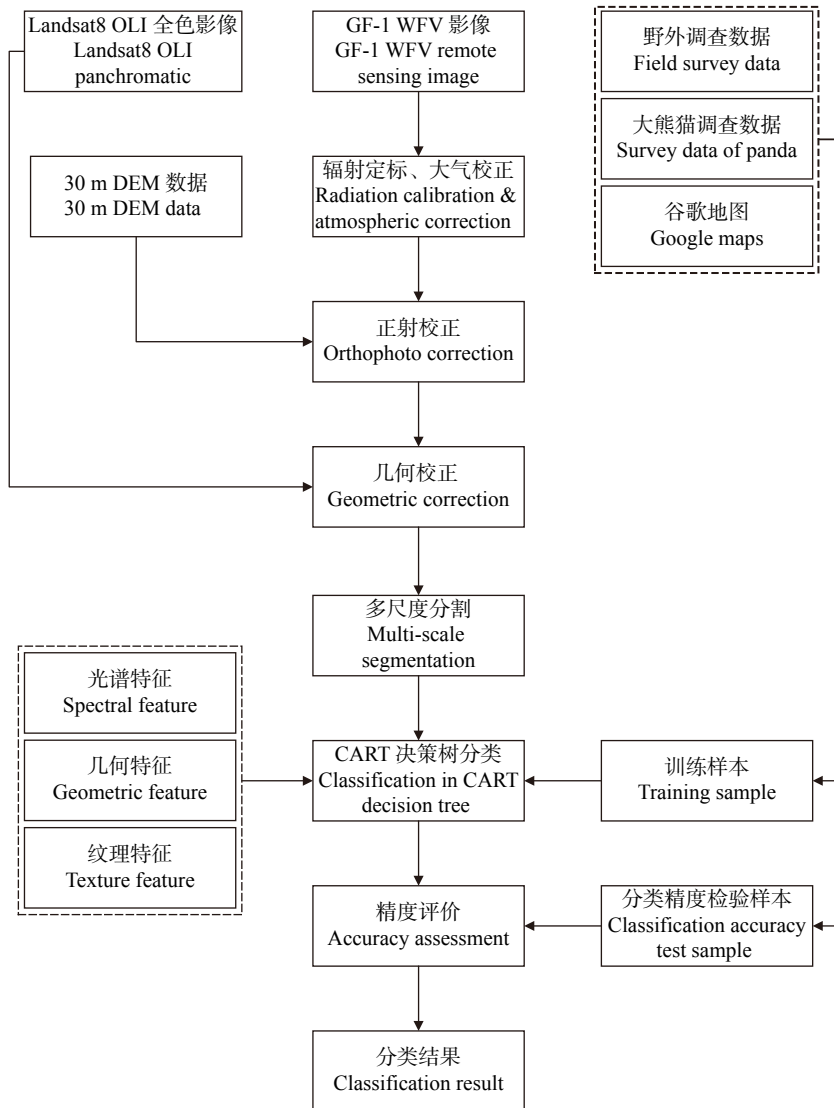


图 2 技术路线图

Fig. 2 Flowchart of the classification technique

和第 4 次大熊猫调查数据确定。将影像分割后的多边形图层与外业调查和第 4 次大熊猫调查的样点数据叠加, 得到对应样点的多边形矢量图层。由于外业调查数据和第 4 次大熊猫调查数据以植被数据为主, 分类系统中耕地、居民区、道路、水体、裸岩样本的选择需结合 Google Earth 进行。综合两种样本选择方法, 各类型样本共 2 220 个, 其中耕地 316 个, 居民区 190 个, 道路 248 个, 水体 190 个, 裸岩 218 个, 针叶林 210 个, 针阔混交林 238 个, 阔叶林 246 个, 灌木林 166 个, 竹林 198 个。各类型样本 50% 用于分类器训练, 50% 用于精度检验。

2.4 特征选择

稳定、有效的特征是提高遥感影像自动解译精度的关键, 依据地表覆盖类别的影像特征, 本研究选择 3 类分类特征。

(1) 光谱特征: 考虑影像对象的像素值及其与其他影像对象像素值之间的关系, 选用亮度(brightness)、

最大差分(max. diff.)、蓝波段(blue)均值、绿波段(green)均值、红波段(red)均值、近红外波段(nir)均值 6 个指标描述影像对象。同时, 考虑研究区地物特征, 添加自定义指数, 其中归一化植被指数 NDVI^[11] 常用于提取植被, 归一化差异水体指数 NDWI^[12] 和归一化差值土壤指数 NDSI 常用于区分水体。研究使用 assign class^[13] 算法区分地物特征与阈值, 选择光谱特征。

$$NDVI = (\rho_{NIR} - \rho_R) / (\rho_{NIR} + \rho_R) \quad (1)$$

$$NDWI = (\rho_G - \rho_{NIR}) / (\rho_G + \rho_{NIR}) \quad (2)$$

$$NDSI = (\rho_R - \rho_B) / (\rho_R + \rho_B) \quad (3)$$

式中: ρ_B 、 ρ_G 、 ρ_R 、 ρ_{NIR} 分别代表 GF-1 WFV 数据在蓝、绿、近红外波段的反射率。

(2) 几何特征: 考虑影像对象的形状和大小, 选用长宽比(length/width)、形状指数(shape index)对

光谱信息相近而形状特征不同的地物进行区分, 尤其是居民区、道路等包含明显几何特征的地物。

(3) 纹理特征: 每种地物都具有特有的纹理结构, 在利用光谱特征和几何特征进行分类的基础上, 通过组合不同灰度共生矩阵(GLCM)的纹理特征进行比较, 选用对比度(contrast)、角二阶矩(ang. 2nd moment)用于描述影像对象的纹理。

2.5 遥感影像分类

决策树是一种直观的知识表示方法, 其以信息论为基础, 将复杂的决策形成过程抽象成易于理解和表达的规则或判断^[14]。CART 分类回归树是 Breiman 提出的一种二分递归分割技术^[15-16], 它将包含测试变量与目标变量构成的训练数据集进行循环迭代分为两个子样本集, 使得生成的每个非叶子节点都有两个分支, 从而形成二叉树形式的决策树结构^[17]。CART 使用的学习样本集如式(4)~(6)所示:

$$L = \{M_1, M_2, \dots, M_m, N\} \quad (4)$$

$$M_1 = (M_{11}, M_{12}, \dots, M_{1t_1}), \dots, M_m = (M_{m1}, M_{m2}, \dots, M_{mt_m}) \quad (5)$$

$$N = (N_1, N_2, \dots, N_k) \quad (6)$$

式中: M_1, M_2, \dots, M_m 为属性向量, N 为标签向量。

CART 选取 GINI 系数最小的属性作为根节点的分裂属性:

$$\text{Gini}(S) = 1 - \sum_{i=1}^n \rho_i^2 \quad (7)$$

式中: 假设 $S = L$ 为样本集, 类别集为 $\{C_1, C_2, \dots, C_n\}$, 共 n 类, 每个类对应 1 个样本子集 $C_i = S_i (1 \leq i \leq n)$ 。令 $|S|$ 为样本集的样本数, $|C_i|$ 为样本集 S 中术语类 C_i 的样本数, 则 $\rho_i = |C_i|/|S|$ 为样本集中样本属于类 C_i 的概率, 通过递归二分成的决策树通常会有“过度拟合”现象, 通过一定的修剪, 最后生成相对最优的二叉树^[18]。

CART 算法基于统计理论进行模式识别, 具有强大的统计解析功能, 相对于二叉树决策树, 它将最优通过树的发现和决策模型的精度验证嵌入到本身的算法之中, 提高规则提取的自动化程度; 同时, 其构建的二叉树结果包含的决策规则更加易懂^[19]。

2.6 精度验证

eCognition 提供 4 种精度检验方法, 即分类稳定性、最佳分类结果统计值、基于 TTA(training and test area)Mask 的混淆矩阵和基于对象样本的混淆矩阵。本研究采用基于 TTA Mask 的混淆矩阵对分类结果进行评价, 分类精度验证采用生产者精度、用户精度、总体精度、Kappa 系数^[20]4 个指标。

$$P_p = P_{kk}/n \quad (8)$$

$$P_u = P_{kk}/j \quad (9)$$

$$P_c = \sum_{k=1}^m P_{kk}/N \quad (10)$$

$$K = \frac{N \sum_{i=1}^m P_{ii} - \sum_{i=1}^m (P_{pi} \times P_{li})}{N^2 - \sum_{i=1}^m (P_{pi} \times P_{li})} \quad (11)$$

式中: P_p 为生产者精度, P_{kk} 为第 k 类的判对样本数, n 为分类器识别为该样本的真实样本数, P_u 为用户精度, j 为地物实际样本数, P_c 为总体精度, m 为分类类别数, N 为样本总数, K 为 Kappa 系数 P_{ii} 为第 i 类的判对样本数, P_{pi} 为某一地物所在列的样本总数, P_{li} 为某一地物所在行的样本总数。

3 结果与分析

3.1 多尺度分割结果

在影像分割过程中分割尺度的选择尤为重要, 分割不足与过度分割都可能导致分类精度的下降, 传统的方法需要根据实际情况反复试验确定最优分割尺度^[21]。最优的分割尺度应该是所分割的结果能通过 1 个或者多个影像对象来表达该地物, 同时又不会出现过于破碎或边界模糊的现象^[22]。本研究影像分割在 eCognition 软件中实现, 分割方法采用多尺度分割。将 GF-1 WFV 影像载入 eCognition 软件中, 4 波段权重均设为 1。分割时, 经多次实践发现, 该研究区形状因子、紧致度分别设置为 0.2 和 0.5 时地物边界显示较好。当形状因子和紧致度固定时, 对比不同分割尺度下的分割结果(图 3), 分割尺度为 30 时, 研究区共划分分割单元 105 219 个, 此时地物较为破碎, 道路缺乏连贯性, 居民区、农田的分割也过于细碎。分割尺度 40 时, 研究区共有分割单元 62 843 个, 此时对象边界与原始数据吻合度较高。当分割尺度为 50 时, 研究区共有分割单元 42 228 个, 此时各类地物的多边形面积增大, 对象边界较模糊, 居民区界限不清, 道路与周边地物也难以分开。对分类结果采用基于样本的混淆矩阵进行精度评

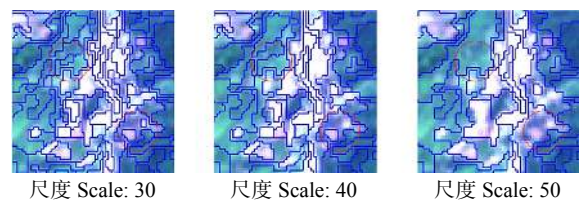


图 3 不同分割尺度分割效果对比

Fig. 3 Comparison of multi-scale segmentation effect with different segmentation scale parameters

价,将分类精度最高的分割尺度确定为最优分割尺度^[23]。由 KIA(Kappa index of agreement)值判断,分类尺度为 30、40 和 50 时, KIA 值分别为 0.740、0.782 和 0.769,故 40 为最优分割尺度。

3.2 影像分类结果

将不同算法的分类结果用 ArcGIS 进行统计(表 2),5 种分类方法对植被与非植被的区分效果都比较好,其中 KNN 和 SVM 分类出的植被面积略大于 CART 决策树。3 类决策树中, CART-1 是采用光谱特征进行决策树分类的结果,林地面积为 1 720.04 km²,其中针阔混交林和阔叶林面积最大,分别为 660.30 km² 和 554.99 km²。采用光谱特征和几何特征构建的决策树分类结果如 CART-2 所示,研究区林地面积 1 718.71 km²,各植被类型面积较 CART-1 分类结果略有变化。CART-3 是组合光谱特征、几何特征和纹理特征构建的决策树,与 CART-2 比较,林地类型中阔叶林面积增加 18.27%,针阔混交林面积减少 14.27%,优势植被类型发生变化。

各植被类型空间分布特征如图 4 所示。针阔混交林和阔叶林是该区主要植被类型,针叶林主要分布在西南高海拔区域,阔叶林分布在中东部地区,灌木林和竹林呈分散分布状态。

3.3 分类精度对比

各分类算法的分类结果混淆矩阵如表 3 至表 7 所示。决策树算法的整体精度均在 83% 以上,较 KNN 算法和 SVM 算法分别提高 7.8% 和 6.89%, Kappa 系数在 0.80 以上,较 KNN 算法的 0.71 和 SVM 算法的 0.72 均有显著提高。3 类决策树算法中,其中基于光谱特征、几何特征和纹理特征的 CART-3 分类效果最好,整体精度为 85.18%, Kappa 系数为

0.83,提取植被信息(针叶林、针阔混交林、阔叶林、灌木林和竹林)的平均生产者精度为 83.42%,平均用户精度为 82.99%。加入几何特征的 CART-2 决策树较 CART-1 决策树整体精度提高 1.10%, Kappa 系数提高 0.01;加入几何特征和纹理特征的 CART-3 决策树较 CART-1 决策树整体精度提高 1.98%, Kappa 系数提高 0.02。在竹林、针叶林和灌木林信息提取时,3 类 CART 决策树精度没有明显差异,但几何特征和纹理特征对提取居民区、道路等地物具有一定作用。

4 结 论

本研究以 GF-1 WFV 影像为数据源,采用面向对象和 CART 决策树的方法对保护区植被类型进行了分类,主要研究结论如下:(1)多尺度分割过程中,分割尺度为 40 时,对象边界与原始地物吻合度较好,此时 KIA 值为 0.782,是研究区最优分割尺度;(2)保护区 90% 以上的地类是林地,针阔混交林、阔叶林和针叶林是本区主要植被类型,占保护区总面积的 76% 以上,其中针叶林主要分布在西南高海拔区域,阔叶林主要分布在中东部地区,灌木林和竹林多呈分散分布;(3)分类精度显示,决策树算法的整体精度均在 83% 以上, Kappa 系数在 0.80 以上,优于 KNN 和 SVM 算法,其中基于光谱特征、几何特征和纹理特征的 CART-3 分类效果最好,整体精度为 85.18%, Kappa 系数为 0.83。(4)基于 CART 决策树算法和面向对象方法的 GF-1 遥感影像分类方法适用于保护区植被类型分布研究,可有效辅助保护区监测工作。

表 2 不同分类算法分类结果统计

Tab. 2 Statistics of classification results in different classification algorithms

类别 Type	CART-1		CART-2		CART-3		KNN		SVM	
	面积 Area/km ²	比例 Ratio/%	面积 Area/km ²	比例 Ratio/%	面积 Area/km ²	比例 Ratio/%	面积 Area/km ²	比例 Ratio/%	面积 Area/km ²	比例 Ratio/%
耕地 Cultivated land	84.08	4.50	88.05	4.71	87.92	4.71	78.71	4.21	73.69	3.95
居民区 Residential area	4.65	0.25	3.97	0.21	4.77	0.26	5.07	0.27	3.57	0.19
道路 Road	24.93	1.33	23.69	1.27	23.07	1.24	11.84	0.63	18.29	0.98
水体 Water area	6.69	0.36	5.52	0.30	5.47	0.29	2.39	0.13	2.81	0.15
裸岩 Bare rock	27.35	1.46	27.81	1.49	27.81	1.49	27.72	1.48	27.34	1.46
针叶林 Coniferous forest	211.41	11.32	211.41	11.32	211.41	11.32	205.83	11.02	208.77	11.18
针阔混交林 Coniferous-broadleaved forest	660.30	35.35	682.27	36.53	584.90	31.32	531.07	28.43	546.93	29.28
阔叶林 Broadleaved forest	554.99	29.71	533.01	28.54	630.38	33.75	723.26	38.72	762.60	40.83
灌木林 Shrubwood	96.08	5.14	94.75	5.07	94.74	5.07	56.11	3.00	71.23	3.81
竹林 Bamboo forest	197.27	10.56	197.27	10.56	197.27	10.56	225.76	12.09	152.52	8.17

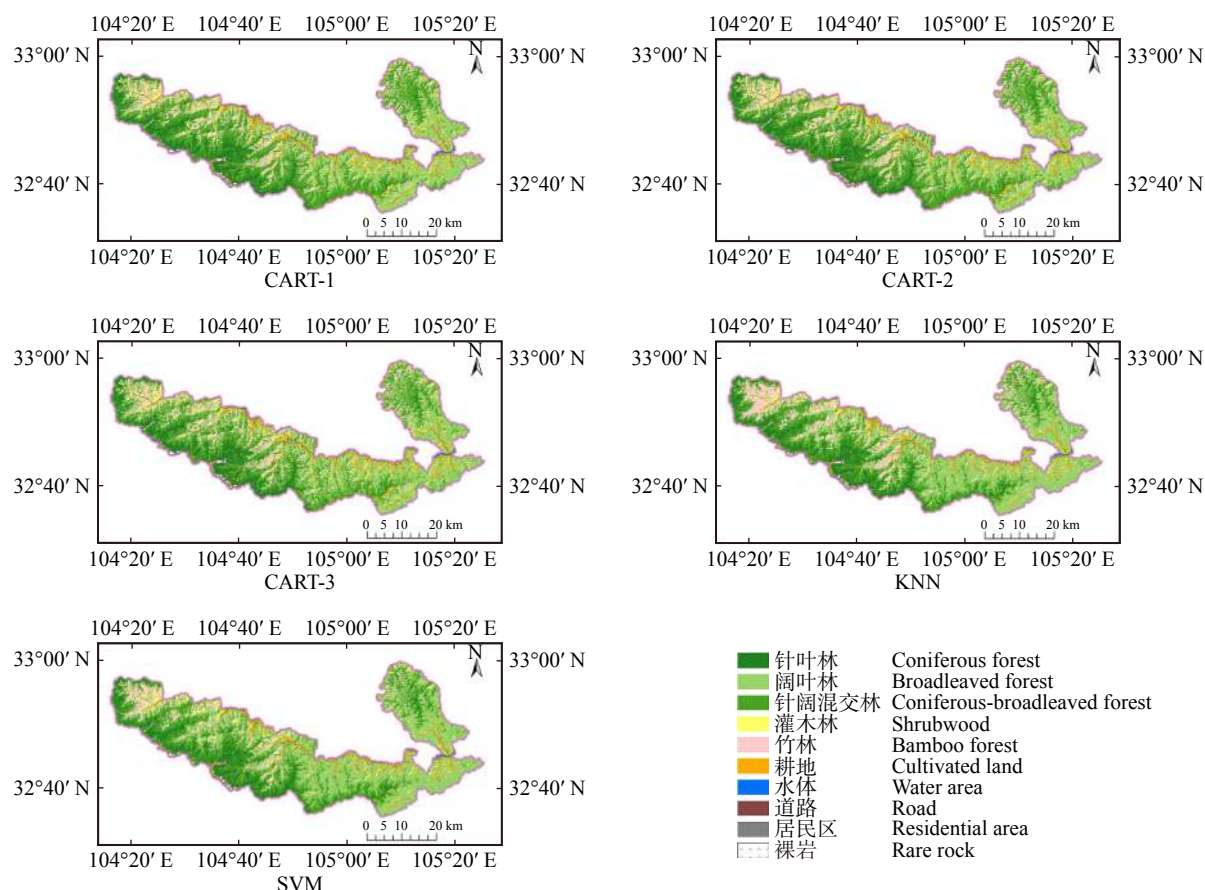


图 4 研究区植被类型分类结果对比图

Fig. 4 Comparison in classification results of vegetation types in study area

表 3 基于光谱特征的决策树分类结果混淆矩阵 (单位: 像元)

Tab. 3 Confusion matrix of CART decision tree classification based on spectral features (unit: pixel)

类别 Type	竹林 Bamboo forest	阔叶林 Broadleaved forest	针阔混交林 Coniferous-broadleaved forest	针叶林 Coniferous forest	灌木林 Shrubwood	居民区 Residential area	道路 Road	耕地 Cultivated land	裸岩 Bare rock	水体 Water area	生产者精度 Producer accuracy/%	用户精度 User accuracy/%
竹林 Bamboo forest	9 896	2 246	1 244	0	1 544	0	0	0	0	0	83.03	66.28
阔叶林 Broadleaved forest	959	15 980	1 458	0	347	0	0	477	0	0	73.74	83.14
针阔混交林 Coniferous-broadleaved forest	617	2 694	16 272	480	176	0	0	0	0	0	80.50	80.40
针叶林 Coniferous forest	0	0	709	18 569	0	0	0	0	0	0	97.48	96.32
灌木林 Shrubwood	375	622	0	0	8 023	0	0	511	0	0	76.10	84.18
居民区 Residential area	0	0	0	0	0	2 190	681	6	461	68	72.02	64.30
道路 Road	0	61	0	0	32	345	6 460	1 624	0	146	81.13	74.53
耕地 Cultivated land	71	68	529	0	395	227	650	11 465	560	0	79.60	82.10
裸岩 Bare rock	0	0	0	0	26	0	79	220	12 124	0	86.38	97.39
水体 Water area	0	0	0	0	0	279	93	100	890	8 335	97.50	85.95

总体精度 Overall accuracy: 83.20%

Kappa系数 Kappa coefficient: 0.80

表4 基于光谱和几何特征的决策树分类结果混淆矩阵(单位:像元)

Tab. 4 Confusion matrix of CART decision tree classification based on spectral and geometric features (unit: pixel)

类别 Type	竹林 Bamboo forest	阔叶林 Broadleaved forest	针阔混交林 Coniferous-broadleaved forest	针叶林 Coniferous forest	灌木林 Shrubwood	居民区 Residential area	道路 Road	耕地 Cultivated land	裸岩 Bare rock	水体 Water area	生产者精度 Producer accuracy/%	用户精度 User accuracy/%
竹林 Bamboo forest	9 896	2 246	1 244	0	1 544	0	0	0	0	0	83.03	66.28
阔叶林 Broadleaved forest	959	15 980	1 035	0	347	0	0	477	0	0	73.74	85.00
针阔混交林 Coniferous-broadleaved forest	617	2 694	16 695	480	176	0	0	0	0	0	82.60	80.80
针叶林 Coniferous forest	0	0	709	18 569	0	0	0	0	0	0	97.48	96.32
灌木林 Shrubwood	375	622	0	0	8 023	0	0	779	0	0	76.10	81.88
居民区 Residential area	0	0	0	0	0	2 001	499	41	224	68	65.80	70.63
道路 Road	0	0	0	0	32	534	7 027	914	374	106	88.25	78.20
耕地 Cultivated land	71	129	529	0	395	227	405	11 872	560	56	82.43	83.35
裸岩 Bare rock	0	0	0	0	26	0	0	272	12 468	102	88.84	96.90
水体 Water area	0	0	0	0	0	279	32	48	409	8 217	96.12	91.45
总体精度 Overall accuracy: 84.30%											Kappa系数 Kappa coefficient: 0.82	

表5 基于光谱、几何和纹理特征决策树分类结果混淆矩阵(单位:像元)

Tab. 5 Confusion matrix of CART decision tree classification based on spectral, geometric and texture features (unit: pixel)

类别 Type	竹林 Bamboo forest	阔叶林 Broadleaved forest	针阔混交林 Coniferous-broadleaved forest	针叶林 Coniferous forest	灌木林 Shrubwood	居民区 Residential area	道路 Road	耕地 Cultivated land	裸岩 Bare rock	水体 Water area	生产者精度 Producer accuracy/%	用户精度 User accuracy/%
竹林 Bamboo forest	9 896	2 246	1 244	0	1 544	0	0	0	0	0	83.03	66.28
阔叶林 Broadleaved forest	959	17 440	1 563	0	347	0	0	477	0	0	80.48	83.90
针阔混交林 Coniferous-broadleaved forest	617	1 234	16 167	480	176	0	0	0	0	0	79.99	86.57
针叶林 Coniferous forest	0	0	709	18 569	0	0	0	0	0	0	97.48	96.32
灌木林 Shrubwood	375	622	0	0	8 023	0	0	779	0	0	76.10	81.88
居民区 Residential area	0	0	0	0	0	2 198	466	41	509	10	72.28	68.18
道路 Road	0	0	0	0	32	363	7 068	914	89	164	88.76	81.90
耕地 Cultivated land	71	129	529	0	395	219	397	11 872	560	56	82.43	83.44
裸岩 Bare rock	0	0	0	0	26	0	0	272	12 468	102	88.84	96.90
水体 Water area	0	0	0	0	0	261	32	48	409	8 217	96.12	91.64
总体精度 Overall accuracy: 85.18%											Kappa系数 Kappa coefficient: 0.83	

表 6 基于最邻近算法的分类结果混淆矩阵 (单位: 像元)

Tab. 6 Confusion matrix based on KNN classification (unit: pixel)

类别 Type	竹林 Bamboo forest	阔叶林 Broadleaved forest	针阔混交林 Coniferous- broadleaved forest	针叶林 Coniferous forest	灌木林 Shrubwood	居民区 Residential area	道路 Road	耕地 Cultivated land	裸岩 Bare rock	水体 Water area	生产者精度 Producer accuracy/%	用户精度 User accuracy/%
竹林 Bamboo forest	6 930	3 914	954	0	3 407	0	0	0	0	0	58.15	45.58
阔叶林 Broadleaved forest	2 397	16 001	2 417	0	222	0	0	1 065	0	0	73.84	72.40
针阔混交林 Coniferous- broadleaved forest	820	708	12 888	480	1 052	0	0	0	0	0	63.76	80.81
针叶林 Coniferous forest	0	0	2 599	18 569	0	0	0	0	0	0	97.48	87.72
灌木林 Shrubwood	1 771	919	1 277	0	5 582	0	0	705	0	0	52.95	54.44
居民区 Residential area	0	0	0	0	0	1 886	633	165	618	605	62.02	48.27
道路 Road	0	0	0	0	0	518	5 277	175	0	101	66.27	86.92
耕地 Cultivated land	0	129	77	0	254	421	1 663	11 312	599	160	78.54	77.40
裸岩 Bare rock	0	0	0	0	26	174	358	981	12 818	150	91.33	88.36
水体 Water area	0	0	0	0	0	42	32	0	0	7 533	88.12	99.03
总体精度 Overall accuracy: 75.20%						Kappa系数 Kappa coefficient: 0.71						

表 7 基于支持向量机算法的分类结果混淆矩阵 (单位: 像元)

Tab. 7 Confusion matrix based on SVM classification (unit: pixel)

类别 Type	竹林 Bamboo forest	阔叶林 Broadleaved forest	针阔混交林 Coniferous- broadleaved forest	针叶林 Coniferous forest	灌木林 Shrubwood	居民区 Residential area	道路 Road	耕地 Cultivated land	裸岩 Bare rock	水体 Water area	生产者精度 Producer accuracy/%	用户精度 User accuracy/%
竹林 Bamboo forest	6 736	2 227	693	0	3 035	0	0	0	0	0	56.52	53.08
阔叶林 Broadleaved forest	2 944	15 996	3 076	0	620	0	0	695	0	0	73.81	78.56
针阔混交林 Coniferous- broadleaved forest	625	1 354	12 669	457	724	0	0	0	0	0	62.68	80.04
针叶林 Coniferous forest	0	0	2 405	18 592	0	0	0	0	0	0	97.60	88.55
灌木林 Shrubwood	1 613	1 719	1 292	0	5 993	0	0	852	0	0	56.84	52.25
居民区 Residential area	0	0	0	0	0	2 061	350	29	618	396	67.77	59.67
道路 Road	0	0	0	0	0	238	6 139	554	41	193	77.10	85.68
耕地 Cultivated land	0	375	77	0	145	545	1 145	10 972	216	128	76.18	80.66
裸岩 Bare rock	0	0	0	0	26	155	278	1 301	13 160	150	93.77	87.33
水体 Water area	0	0	0	0	0	42	51	0	0	7 682	89.86	98.80
总体精度 Overall accuracy: 76.11%						Kappa系数 Kappa coefficient: 0.72						

参 考 文 献

- [1] Howard P C, Davenport T R B, Kigenyi F W, et al. Protected area planning in the tropics: Uganda's national system of forest nature reserves[J]. *Conservation Biology*, 2000, 14(3): 858–875.
- [2] Radeloff V C, Stewart S I, Hawbaker T J, et al. Housing growth in and near United States protected areas limits their conservation value[J]. *Proceedings of the National Academy of Sciences USA*, 2010, 107(2): 940–945.
- [3] Stein B A, Scott C, Benton N. Federal lands and endangered species: the role of military and other federal lands in sustaining biodiversity[J]. *Bioscience*, 2008, 58: 339–347.
- [4] Millennium Ecosystem Assessment Board. Ecosystems and human well-being[M]. Washington: Island Press, 2003.
- [5] Maiorano L, Falucci A, Boitani L. Size-dependent resistance of protected areas to land-use change[J]. *Proceedings of the Royal Society B: Biological Sciences*, 2008, 275: 1297–1304.
- [6] Heenkenda M K, Joyce K E, Maier S W, et al. Mangrove species identification: comparing WorldView-2 with aerial photographs[J]. *Remote Sensing*, 2014, 6(7): 6064–6088.
- [7] Sun Z C, Leinenkugel P, Guo H D, et al. Extracting distribution and expansion of rubber plantations from Landsat imagery using the C5.0 decision tree method[J/OL]. *Journal of Applied Remote Sensing*, 2017, 11(2): 026011[2019-03-11]. <http://doi:10.1117/1.JRS.11.026011>.
- [8] 郝泷, 陈永富, 刘华, 等. 基于纹理信息 CART 决策树的林芝县森林植被面向对象分类[J]. *遥感技术与应用*, 2017, 32(2): 386–394.
- Hao S, Chen Y F, Liu H, et al. Object-oriented forest classification of Linzhi County based on CART decision tree with texture information[J]. *Remote Sensing Technology and Application*, 2017, 32(2): 386–394.
- [9] 雷光斌, 李爱农, 谭剑波, 等. 基于多源多时相遥感影像的山地森林分类决策树模型研究[J]. *遥感技术与应用*, 2016, 31(1): 31–41.
- Lei G B, Li A N, Tan J B, et al. Forest type mapping in mountainous area using multisource and multi-temporal satellite images and decision tree models[J]. *Remote Sensing Technology and Application*, 2016, 31(1): 31–41.
- [10] 钱军朝, 徐丽华, 邱布布, 等. 基于 WorldView-2 影像数据对杭州西湖绿地信息提取研究[J]. *西南林业大学学报*, 2017, 37(4): 156–166.
- Qian J C, Xu L H, Qiu B B, et al. Extraction of urban green space based on WorldView-2 images in West Lake District of Hangzhou[J]. *Journal of Southwest Forestry University*, 2017, 37(4): 156–166.
- [11] Rouse J W J, Haas R H, Schell J A, et al. Monitoring vegetation systems in the Great Plains with ERTS[J]. *Nasa Special Publication*, 1974, 351: 309–314.
- [12] Gao B C. NDWI: a normalized difference water index for remote sensing of vegetation liquid water from space[J]. *Remote Sensing of Environment*, 1996, 58(3): 357–366.
- [13] 蔡耀君, 华璿, 卢远. 峰丛洼地农作物面向对象信息提取规则集[J]. *遥感信息*, 2018, 34(10): 245–252.
- Cai Y J, Hua C, Lu Y. Rules for crop extraction on peak cluster of karst area using object-oriented classification technology[J]. *Remote Sensing Information*, 2018, 34(10): 245–252.
- [14] 潘琛, 杜培军, 张海荣. 决策树分类算法及其在遥感图像处理中的应用[J]. *测绘科学*, 2008, 33(1): 208–211.
- Pan C, Du P J, Zhang H R. Decision tree classification and its application in processing of remote sensing images[J]. *Science of Surveying and Mapping*, 2008, 33(1): 208–211.
- [15] Breiman L, Friedman J B, Stone C J, et al. Classification and regression tree[M]. Boca Raton FL: Chapman & Hall/CRC, 1984.
- [16] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- Zhou Z H. Machine learning[M]. Beijing: Tsinghua University Press, 2016.
- [17] 陈云, 戴锦芳, 李俊杰. 基于影像多种特征的 CART 决策树分类方法及应用[J]. *地理与地理信息科学*, 2008, 24(2): 33–36.
- Chen Y, Dai J F, Li J J. CART-based decision tree classifier using multi-feature of image and its application[J]. *Geography and Geo-Information of Science*, 2008, 24(2): 33–36.
- [18] 马宇龙, 林志垒. 基于面向对象和 CART 决策树方法的遥感影像湿地变化检测研究: 以龙祥岛地区为例[J]. *福建师范大学学报(自然科学版)*, 2017, 33(6): 69–80.
- Ma Y L, Dai Z L. Wetland change detection based on object-oriented and CART decision tree method[J]. *Journal of Fujian Normal University (Natural Science Edition)*, 2017, 33(6): 69–80.
- [19] 孙建伟, 王超, 王娜, 等. 基于 CART 决策树的 ZY-3 卫星遥感数据土地利用分类监测[J]. *华中师范大学学报(自然科学版)*, 2016, 50(6): 937–943.
- Sun J W, Wang C, Wang N, et al. Research on land use classification monitoring through the remote sensing data of ZY-3 satellite based on CART decision tree[J]. *Journal of Central China Normal University (Natural Science Edition)*, 2016, 50(6): 937–943.
- [20] 张小平, 曹卫彬, 刘姣娣. 基于遥感影像的棉花种植面积提取方法研究[J]. *安徽农业科学*, 2011, 39(7): 4226–4228, 4297.
- Zhang X P, Cao W B, Liu J D. Study on the extraction method for cotton-growing area based on remote sensing image[J]. *Journal of Anhui Agriculture Science*, 2011, 39(7): 4226–4228, 4297.
- [21] 陈天博, 胡卓玮, 魏隼, 等. 无人机遥感数据处理与滑坡信息提取[J]. *地球信息科学*, 2017, 19(5): 692–701.
- Chen T B, Hu Z W, Wei L, et al. Data processing and landslide information extraction based on UAV remote sensing[J]. *Journal of Geo-information Science*, 2017, 19(5): 692–701.
- [22] 李晓红, 陈尔学, 李增元, 等. 综合应用多源遥感数据的面向对象土地覆盖分类方法[J]. *林业科学*, 2018, 54(2): 68–80.
- Li X H, Chen E X, Li Z Y, et al. Object based land cover classification method integrating multi-source remote sensing data[J]. *Scientia Silvae Sinicae*, 2018, 54(2): 68–80.
- [23] 魏晶昱, 毛学刚, 方本煜, 等. 基于 Landsat 8 OLI 辅助的亚米级遥感影像树种识别[J]. *北京林业大学学报*, 2016, 38(11): 23–33.
- Wei J Y, Mao X G, Fang B Y, et al. Submeter remote sensing image recognition of trees based on Landsat 8 OLI support[J]. *Journal of Beijing Forestry University*, 2016, 38(11): 23–33.

(责任编辑 范娟
责任编委 崔国发)