

引用格式: Li Meng, Nian Yanyun, Bian Rui, et al. Classification of *Picea crassifolia* and *Sabina przewalskii* based on Multi-source Remote Sensing Images[J]. Remote Sensing Technology and Application, 2020, 35(4): 855-863. [李萌, 年雁云, 边瑞, 等. 基于多源遥感影像的青海云杉和祁连圆柏分类[J]. 遥感技术与应用, 2020, 35(4): 855-863.] doi: 10.11873/j.issn.1004-0323.2020.4.0855

基于多源遥感影像的青海云杉和祁连圆柏分类

李 萌, 年雁云, 边 瑞, 白艳萍, 马金辉

(兰州大学资源环境学院, 甘肃 兰州 730000)

摘要: 青海云杉和祁连圆柏是祁连山自然保护区的优势种, 提取两种类型树木的空间分布对保护区森林资源的管理和监测方面有重要意义。使用 Sentinel-2A(S2)、Sentinel-1A(S1)、Landsat-8(L8) 3种遥感影像及来自 SRTM DEM 的地形数据, 基于随机森林分类方法, 设置 8 种组合方案共 22 个特征变量, 以祁连山东段的甘肃连城自然保护区为例, 对青海云杉和祁连圆柏进行分类试验。结果表明: 融合 Sentinel-1A(S1) 数据的 VV 和 VH 两种后向散射信息的精度最高, 达到 92.85%, 比使用单一影像 Landsat-8 提高了 11.64%。实验表明: 结合多源遥感影像的不同波段信息是提高森林类型分类精度的有效手段, 有助于复杂山区森林资源调查、植被信息提取等需求。

关键词: Sentinel-2A; 特征变量; 随机森林; 信息提取; 连城自然保护区

中图分类号: S757; TP79 **文献标志码:** A **文章编号:** 1004-0323(2020)04-0855-09

1 引言

森林作为陆地生态系统的主体, 在减缓气候变化, 改善环境及保护生态安全等方面有重要意义。传统的森林类型识别以野外调查为主, 成本高、效率低, 周期长, 时效性和准确难以满足林业发展的需求^[1]。而遥感技术具有成本低、可重复、覆盖范围大等优势, 将其应用到森林类型的识别和森林资源信息中具有重要的意义。由于遥感技术可用于地形复杂或难以到达的地区获取森林信息, 与传统方法互补, 减少实地工作量, 所以有必要探索多源遥感数据获取森林类型信息的潜力。

近年来, 对于森林类型识别已有大量相关研究, 主要包括森林类型分类和树种识别^[2-4]。李梦颖等^[5]利用 Landsat-8 影像对某天然林区森林类型分类, 结果表明结合纹理特征与光谱特征能够提高分类精度。田甜等^[6]从 QuickBird 数据提取纹理特征、光谱特征、植被指数对福建省将乐林场的优势树种类型提取。Markus 等^[7]借助 Sentinel-2 影像分别用

基于对象和基于像元两种方法对德国巴伐利亚州两个林区内 7 种树种进行分类。Zhu 等^[8]使用多时相 Landsat TM 影像, 利用支持向量机方法对阔叶林进行精细分类。Xie 等^[9]使用资源 3 号影像对某林场的 4 种针叶林和 3 种落叶阔叶林进行分类, 结果表明树冠高度特征可以改善一些树种的分类精度。以上研究大多使用单一数据源的光学影像, 而近年遥感数据源向着全天候、多分辨率、多传感器等方向快速发展, 相关研究认为结合多源遥感数据, 充分利用不同数据的波段信息, 优势互补, 能够更好地识别地物, 已有学者将多源遥感数据组合的方法应用到森林资源分类与识别当中^[10-13]。雷光斌等^[14]利用 Landsat TM 和 CCD 影像对岷江上游地区森林类型分类, 总体精度达到 87.05%。刘刚等^[15]结合 SPOT5、Landsat TM、CBERS 3 种遥感数据对某林场进行森林信息提取, 结果表明多源融合数据使各类型分类精度都有所提高。虽然使用多源遥感影像进行森林信息提取的研究较多, 但将合成孔径

收稿日期: 2019-08-29; 修订日期: 2020-07-13

基金项目: 中国科学院 A 类战略性先导科技专项(XDA20100102)资助。

作者简介: 李 萌(1994—), 女, 甘肃庆阳人, 硕士研究生, 主要从事遥感与 GIS 应用研究。E-mail: limeng17@lzu.edu.cn

通讯作者: 年雁云(1979—), 男, 甘肃临夏人, 副教授, 主要从事无人机遥感和 GIS 应用研究。E-mail: yynian@lzu.edu.cn

雷达(SAR, Synthetic Aperture Radar)数据引入到森林分类的研究较为缺乏,尤其针对地形复杂、云雾雨雪较多、数据获取困难的山地森林,仅用单一数据源难以满足分类精度需求。SAR 卫星不受天气因素的影响,全天候、全天时不间断工作,能够获取云、雨、雾影响下的地表信息, Sentinel-2A 影像具有覆盖宽、多分辨率、丰富的光谱特征等优势,本文尝试结合光学影像和 SAR 数据,探讨多源遥感影像在山地森林信息提取的贡献和作用,促进 SAR 数据在森林类型分类研究中的应用。

在不同数据源的森林类型识别研究中,众多分类方法的选择也很关键。其中,基于多棵 CART 决策树构成的随机森林算法在训练样本不足的情况下也能达到较高的精度^[16],已有部分学者将其应用到森林资源分类研究当中,并取得较好的分类结果^[17-20]。因此本文考虑利用 Sentinel-1A、Sentinel-2A, Landsat-8 影像及 DEM 数据,基于随机森林分类算法,对甘肃连城国家级自然保护区内青海云杉和祁连圆柏进行分类试验,并分析不同特征变量的重要性。

2 研究区概况、数据源及预处理

2.1 研究区概况

甘肃连城国家级自然保护区,处于祁连山山地与黄土高原的重要过渡区域(范围大致覆盖了 102°26′~102°55′ E, 36°33′~36°48′ N),且位于黄河上游的重要支流——大通河中下游,其地理位置特殊,森林生态系统完整,是开展半干旱地区山地森林研究的典型区域(图 1)。研究区内主要地物类型为青海云杉、祁连圆柏、高山灌丛草甸、山地森林草原、裸岩等,特殊的地理位置和地形地貌使保护区境内自然条件复杂多样,为森林资源的多样性提供了多样的生存空间。

2.2 数据源及预处理

2.2.1 多光谱数据

研究使用的多光谱数据(表 1、表 2)是 Sentinel-2A 和 Landsat-8(OLI)影像。Sentinel-2 影像具有 10 m、20 m、60 m 3 种分辨率,包括可见光、近红外、植被红边等共 13 个波段,研究所用影像从欧空局数据 中 心 (<https://scihub. Copernicus. eu/dhus/#/home>)下载,选择覆盖研究区无云且质量良好(2018 年 7 月 17 日)的两幅等级为 1C 的 Sentinel-2A 影像,该数据是经过正射校正和亚像元级几何精校正的

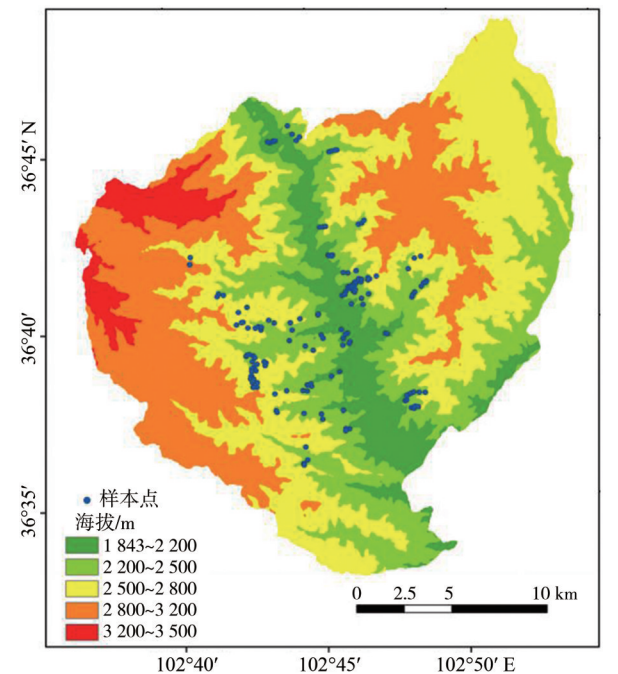


图 1 研究区位置及样本点分布示意

Fig.1 Location of the study area and distribution of sample sites

表 1 3 种遥感数据的基本信息

Table 1 Basic information of three types of remote sensing data

数据	重访周期/d	分辨率/m	波段数或波段名
Sentinel-2A	10	10/20/60	13
Landsat-8	16	30/15	11
Sentinel-1A	12	5×20	C 波段

表 2 Sentinel-2A 波段信息

Table 2 Band information of Sentinel-2A

分辨率/m	波段名称	中心波长/nm	编号
10	Blue	490	b2
	Green	560	b3
	Red	665	b4
	NIR	842	b8
20	Vegetation Red Edge1	705	b5
	Vegetation Red Edge2	740	b6
	Vegetation Red Edge3	783	b7
	Narrow NIR	865	b8a
	SWIR 1	1610	b11
	SWIR 2	2190	b12
60	Coastal aerosol	443	b1
	Water vapour	940	b9
	SWIR_cirrus	1375	b10

大气表观反射率产品,因此只需进行大气校正。本文使用 Sentinel-2A 分辨率为 10 m 和 20 m 的波段。借助欧洲航空局 ESA 提供的 Sen2cor 插件对 Senti-

nel-2A影像进行大气校正,再使用SNAP软件中的最近邻法将分辨率为20 m的6个波段重采样到10 m,并结合4个原始10 m分辨率的波段共得到10个10 m分辨率的波段,转换成ENVI标准格式,在ENVI5.3里进行影像镶嵌和裁剪。

Landsat-8(OLI)影像有9个波段,本实验使用前7个30 m分辨率的波段。从美国地质调查局网站(<https://glovis.usgs.gov/>)获取一幅研究区范围内无云的Landsat-8影像(2018年5月16日),利用ENVI5.3软件对影像进行辐射定标、大气校正、图像裁剪等处理,将分辨率重采样到10 m。

2.2.2 合成孔径雷达(SAR)卫星数据

Sentinel-1卫星是全球环境与安全监测计划(Global Monitoring for Environment and Security)中的地球观测卫星,由A、B两颗极轨卫星组成,载有C波段合成孔径雷达,全天候、全天时不间断运行,无论天气状况如何,都能获取云覆盖下的连续图像。从欧空局数据中心选取干涉测量宽幅模式(IW)的双极化(VV+VH)的Level-1级GRD格式的Sentinel-1A雷达数据(2018年9月23日)。在SARscape 5.2.1中对获取的研究区Sentinel-1A原始数据进行地理编码、辐射定标,使用Frost滤波器进行滤波以减少斑点噪声,为避免小值影响,将处理结果转换为以dB为单位的后向散射系数图像,如下所示:

$$DN_{dB} = 10 \cdot \log_{10}(N) \quad (1)$$

其中: N 是通过地理编码和辐射定标计算的像素值。利用ENVI5.3对所得的后向散射系数图像进行研究区范围裁剪,并重采样到10 m分辨率。

2.2.3 SRTM DEM数据

SRTM DEM(Shuttle Radar Topography Mission DEM)数据是由美国国家航空航天局(NASA, National Aeronautics and Space Administration)和美国国家地理空间情报局(NGIA, National Geospatial-Intelligence Agency)合作采集和发布的数字高程模型,包括SRTM1(30 m)和SRTM3(90 m),该类型数据覆盖范围广、可免费下载,在地学研究领域应用广泛。获取覆盖研究区范围30 m分辨率的SRTM1 DEM,利用ENVI5.3和ArcGIS10.2进行裁剪,提取坡度、坡向,并重采样到10 m分辨率。

2.2.4 样本数据

使用的样本数据来源于第二次青藏高原综合科学考察研究于2018年9月28日到10月10日野外

实地考察,借助无人机与相机拍摄记录青海云杉和祁连圆柏以及其他林地的树木形状特征、分布特征,以此辅助Google Earth高分辨率遥感影像目视解译(图2)。通过以上方式获得保护区内样本474个,用于分类模型建立和精度验证,其中青海云杉80个,祁连圆柏82个,其他林地101个,非林地211个。研究区的林地大部分位于该区中部,而青海云杉和祁连圆柏多分布林区中央、河谷地带,因此如图2所示,青海云杉和祁连圆柏的样本点分布较为集中在整个研究区中间部分。

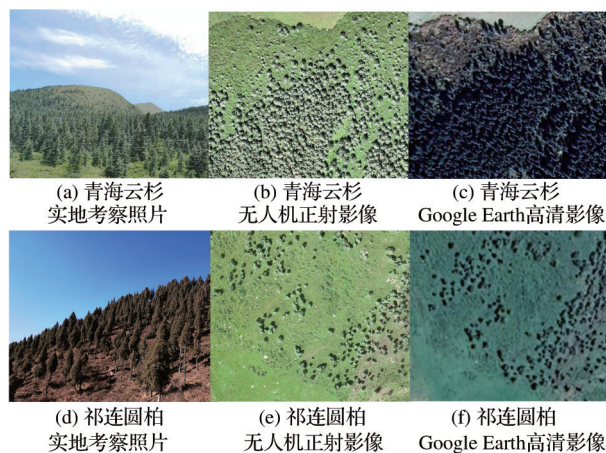


图2 优势树种不同图像对比

Fig.2 Comparison of dominant tree species images

表3 优势树种解译特征对比

Table 3 Comparison of interpretation characteristics of dominant tree species

优势树种	解译特征描述
青海云杉	坡向: 主要分布在阴坡、半阴坡和河谷地带;
	形状: 顶部比较尖, 树形又高又尖;
	分布: 呈大片密集状态紧凑分布;
	色调: 偏暗偏深
祁连圆柏	坡向: 主要分布在阳坡;
	形状: 树高较低、顶部圆润, 类似棒槌形状;
	分布: 呈零星状分布;
	色调: 较亮

3 研究方法

3.1 遥感特征选取

研究区属于干旱半干旱山地森林区域,其森林结构复杂,青海云杉和祁连圆柏的生境特点明显,两者的空间分布在坡向上有明显差异,青海云杉主要分布在阴坡和半阴坡,而圆柏主要分布在阳坡。图3是青海云杉和祁连圆柏在Landsat-8和Sentinel-2A影像中的平均光谱特征,可以看出Sentinel-

2A中可见光波段光谱差异比Landsat-8明显,在植被红边波段和近红外波段,Sentinel-2A影像对云杉和圆柏的区分能力优于Landsat-8,总体来看,Landsat-8和Sentinel-2A影像中青海云杉与祁连圆柏的光谱差异在近红外波段、短波红外波段及植被红边

波段都大于可见光波段。综合考虑云杉和圆柏在空间分布、光谱方面存在的以上差异以及两种树木在树高、冠幅的差别,选取光谱特征、植被指数、地形因子、纹理特征及后向散射系数共计22个特征变量,作为随机森林分类的输入量。

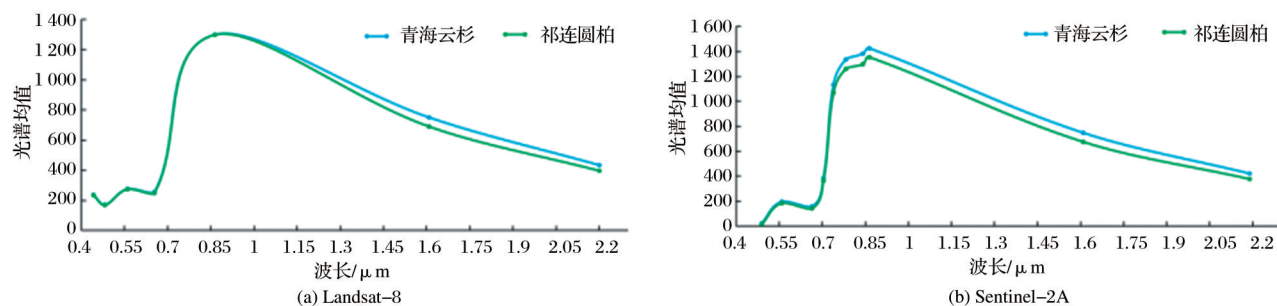


图3 Landsat-8和Sentinel-2A影像中青海云杉和祁连圆柏的平均光谱特征

Fig.3 The mean spectral characteristic of *Picea crassifolia* and *Sabina przewalskii* in Landsat-8 and Sentinel-2A imagery

3.2 随机森林算法

随机森林算法由Breiman于2001年提出,是一种利用多个分类树对数据进行判别与分类的方法,该方法基本原理是利用bootstrap重抽样方法从原始样本选取多个子样本,并逐一进行决策树建模,每棵树单独完成分类后,由投票得出最终输出的分类结果^[14]。本文基于EnMAP2.2.1软件实现随机森林算法的分类,设置提取林地的4种组合方式和优势树种分类的8种组合方式,依次输入包含不同特征

变量的组合以及训练样本集,设置决策树数量为100,每个节点的特征数为默认的特征总数的平方根,计算不同组合方式下RFC(Random Forest Classifier)模型,提取林地和优势树种,获取分类精度最高的一组特征组合。

3.3 分层分类

森林优势树种识别是通过分层分类方法获得的,包括第一层的植被提取,第二层的林地提取和第三层的优势树种分类,技术路线如图4。

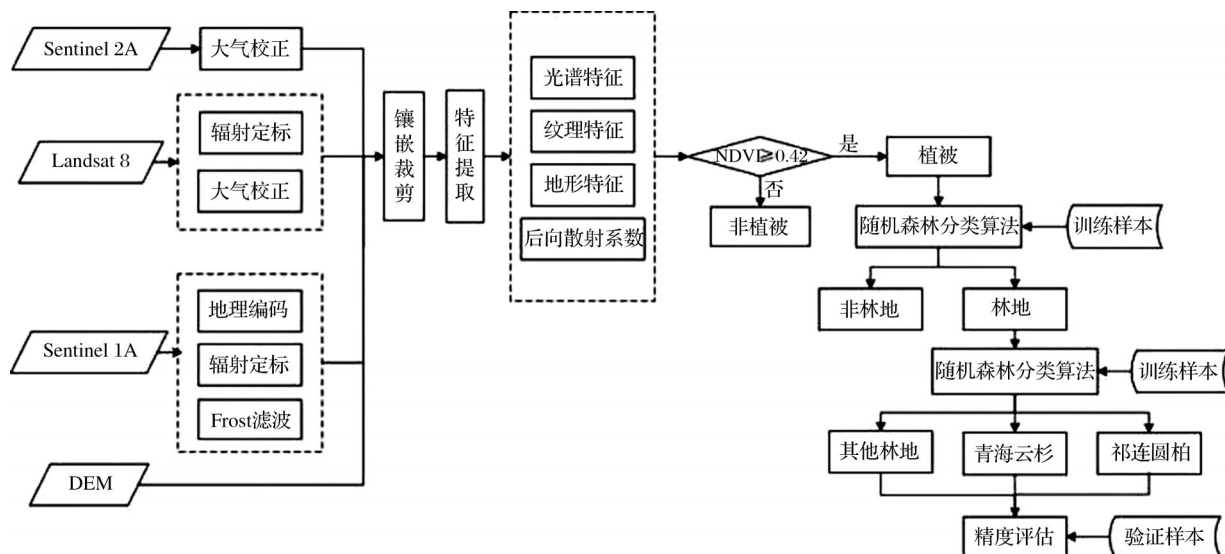


图4 技术流程

Fig.4 Methodological flowchart

3.3.1 第一层植被提取

第一层分类先从预处理的影像提取植被,使用归一化植被指数(NDVI)进行植被提取,再通过比率蓝波段指数去除错误分类为植被的蓝色屋顶。

由于研究区河谷地带的居民区存在一些蓝色屋顶的彩钢房,这种材料的屋顶光谱反射特点与植被相似,值与NDVI相近,导致NDVI提取过程中无法提取纯净的植被,因此需要利用比率蓝波段指数

(RBI)去除错误提取为植被的蓝色屋顶。

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \quad (2)$$

其中: ρ_{nir} 和 ρ_{red} 分别是Sentinel-2A影像的近红外波段和红光波段的反射率。首先使用逐步逼近法进行NDVI的阈值分析,该方法的思想是先在NDVI直方图中确定初始阈值,然后通过目视检查确定植被与影像最佳匹配时的阈值,再利用逐步逼近法对比率蓝波段指数(RBI)进行阈值分析。

$$RBI = \frac{\rho_{nir}}{\rho_{blue}} \quad (3)$$

其中: ρ_{nir} 和 ρ_{blue} 分别是Sentinel-2A影像的近红外波段和蓝光波段的反射率。经多次试验后,设定非植被区移除的NDVI阈值为0.42,蓝色屋顶清除的RBI阈值为2.3,最终提取出纯净的植被层。

3.3.2 第二层林地提取

第二层林地从植被层提取,通过四次实验评估S2、S1、DEM和L8对林地提取的相对贡献,包括光谱特征、植被指数、纹理特征和地形特征,用这些特征(表4)进行随机森林运算以提取林地,输出林地区域作为第三层优势种分类的掩膜。

表4 用于第二层分类的特征变量

Table 4 Feature variables derived from data for the Level-2 classification

数据	特征类型	特征名称	描述
Sentinel-2A 光谱 (10 bands)	特征	S2mean	波段 2-8、8a、11-12 的均值
		S2rvi	NIR/R
		S2evi	$2.5 \times ((NIR-R)/(NIR+6 \times R-7.5 \times B+1))$
		S2dvi	NIR-R
		S2ndvi	$(NIR-R)/(NIR+R)$
		S2b78a	$(NIR2-RE3)/(NIR2+RE3)$
		S2b67	$(RE3-RE2)/(RE3+RE2)$
		S2b58a	$(NIR2-RE1)/(NIR2+RE1)$
		S2b56	$(RE2-RE1)/(RE2+RE1)$
		S2b57	$(RE2-RE1)/(RE2+RE1)$
Sentinel-2A 纹理 (4 bands)	特征	S2b68a	$(NIR2-RE2)/(NIR2+RE2)$
		S2b48a	$(NIR2-R)/(NIR2+R)$
		PCA1	B、G、R、NIR 基于 PCA 分析的第一主成分
		Elevation	DEM 的高程值
DEM	地形特征	Aspect	基于 DEM 提取的坡向
		Slope	基于 DEM 提取的坡度

注:S2表示Sentinel-2A,b表示band,数字代表参与运算的波段编号,具体编号见表2。NIR表示近红外波段,R表示红光波段,G表示绿光波段,B表示蓝光波段,NIR2表示窄近红外波段,RE1、RE2、RE3分别代表3个植被红边波段。

3.3.3 第三层优势种分类

第三层分类是区分林地区域的优势种,获取青海云杉和祁连圆柏的空间分布状况。在光谱特征和地形特征基础上添加后向散射特征以改善分类结果,包括归一化植被指数(NDVI)、比值植被指数(RVI)、增强型植被指数(EVI)和差值植被指数(DVI)和后向散射系数等22个特征变量(表5),以林地作为掩膜基于随机森林算法对青海云杉和祁连圆柏分类。

表5 用于第三层分类的特征变量

Table 5 Feature variables derived from data for the Level-3 classification

数据	特征名称	描述
Sentinel-2A (10 bands)	S2mean	波段 2-8、8a、11-12 的均值
	S2rvi	NIR/R
	S2evi	$2.5 \times ((NIR-R)/(NIR+6 \times R-7.5 \times B+1))$
	S2dvi	NIR-R
	S2ndvi	$(NIR-R)/(NIR+R)$
	S2b78a	$(NIR2-RE3)/(NIR2+RE3)$
	S2b67	$(RE3-RE2)/(RE3+RE2)$
	S2b58a	$(NIR2-RE1)/(NIR2+RE1)$
	S2b56	$(RE2-RE1)/(RE2+RE1)$
	S2b57	$(RE2-RE1)/(RE2+RE1)$
Landsat-8 (7 bands)	S2b68a	$(NIR2-RE2)/(NIR2+RE2)$
	S2b48a	$(NIR2-R)/(NIR2+R)$
	L8mean	波段 1-7 的均值
	L8ndvi	$(NIR-R)/(NIR+R)$
DEM	L8rvi	NIR/R
	L8evi	$2.5 \times ((NIR-R)/(NIR+6 \times R-7.5 \times B+1))$
	L8dvi	NIR-R
DEM	Elevation	DEM 的高程值
	Aspect	基于 DEM 提取的坡向
	Slope	基于 DEM 提取的坡度
Sentinel-1	VV	VV 极化方式的后向散射系数
	VH	VH 极化方式的后向散射系数

注:S2表示Sentinel-2A,L8表示Landsat-8,b表示band,数字代表参与运算的波段编号,具体编号见表2。NIR表示近红外波段,R表示红光波段,G表示绿光波段,B表示蓝光波段,NIR2表示窄近红外波段,RE1、RE2、RE3分别代表3个植被红边波段。

4 结果与分析

4.1 分类结果与精度评价

Sentinel-2A数据和地形特征组合提取林地效果最佳,总体精度为98.43%,Kappa系数为0.9312,提取结果如图5所示。研究区内地物类型多样,植被类型丰富,使用分层提取能提高分类精度。第一层植被提在去除河流、裸岩、建筑用地时也将河谷

地区与植被易于混淆的一小部分蓝色厂房屋顶移除。在纯净的植被层中提取第二层林地时去除高山草甸、灌丛、耕地。为了能准确提取山区林地范围,结合地形特征与光谱特征、纹理特征,对训练样本和验证样本比例设定多次实验,获得取了该区域林地范围。林地范围作为优势树种分类的掩膜,其精度影响优势树种的精度,表 6 为 4 种组合的林地提取精度,选取精度最高为 98.43% 的林地提取结果作为下一步对青海云杉和祁连圆柏进行分类的基础,比较多源遥感数据的 8 种组合的优势种分类精度。

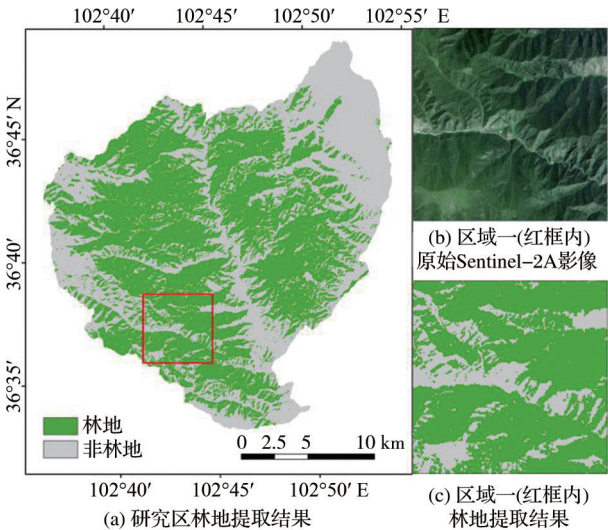


图 5 研究区林地提取结果及细节对比

Fig.5 The extraction result of forestland in study area and details comparison

表 6 4 种组合的林地提取精度

Table 6 Extraction accuracy of forestland for the four types of combinations

组合	数据	总体精度/%	Kappa 系数
1	S2	96.91	0.9312
2	S2+DEM	98.43	0.9681
3	S2+L8	97.48	0.9486
4	S2+S1	97.47	0.9487

对比分析 8 种组合的优势树种分类精度(表 7)可知,组合 1 和 2 都仅使用了光谱特征,但因 Sentinel-2A 比 Landsat 8 光谱信息丰富, Sentinel2A 具有的 3 个植被红边波段和两个近红外波段有利于森林类型的识别,且 Sentinel-2A 比 Landsat-8 分辨率更高,使得组合 2 比 1 获得更高的精度。选取 Sentinel-2A 数据为基础,逐步添加 Landsat-8 数据的光谱特征,DEM 数据的地形特征, Sentinel-1A 数据的

表 7 8 种组合的优势种分类精度

Table 7 Classification accuracy of dominant tree species for the eight types of combinations

组合	数据或特征变量	总体精度 /%	Kappa 系数
1	L8	81.21	0.7240
2	S2	88.04	0.8241
3	S2+Elevation+ Aspect+Slope	90.54	0.8608
4	S2+L8	89.59	0.8485
5	S2+L8+Elevation+ Aspect+Slope	92.83	0.8957
6	S2+L8+Elevation+ Aspect+Slope+ VV	91.24	0.8724
7	S2+L8+Elevation+ Aspect+Slope+ VH	91.23	0.8722
8	S2+L8+Elevation+ Aspect+Slope+ VV+ VH	92.85	0.8958

后向散射信息进一步计算并分析。组合 4 结合两种光学影像得到仅使用光谱信息组合中精度最高的分类结果,比使用单一数据源的组合 1 和 2 分别提高 8.38% 和 1.55%。与仅使用 Sentinel-2A 数据相比,当结合 DEM 数据,总体精度提高 2.50%,该结果也证明了青海云杉和祁连圆柏的空间分布受地形因素的影响较大,而事实两者在海拔和坡向方面分布差异明显。组合 6 和 7 在包含所有光谱特征和地形特征的基础上分别添加 VV 和 VH 后向散射系数,总体精度并没有提高反而稍有降低,且两组的分类精度差异并不明显。当组合 8 使用光谱、地形、后向散射特征共 22 个变量时分类精度达到最高,总体精度为 92.85%, Kappa 系数为 0.895 8,比仅使用单一遥感数据源的组合 1 和 2 分别提高 11.64% 和 4.81%。

组合 8 中青海云杉和祁连圆柏的分类精度如表 8 所示。青海云杉和祁连圆柏是该区域内的优势种,祁连圆柏又是该区域的建群种之一,本文重点关注青海云杉和祁连圆柏的空间分布状况,因此将研究区内油松、落叶松、山杨、白桦等类型归为其他林地,分类结果(图 6)可以清晰看出青海云杉和祁连圆柏的整体分布状况和坡向上的分布差异,两者空间分布趋势与祁连山自然保护区所绘制的连城国家级自然保护区植被图中云杉林和圆柏林一致。

4.2 特征重要性评价

随机森林算法在对输入的特征变量进行分类

表 8 组合 8 的青海云杉和祁连圆柏的分类精度

Table 8 Classification accuracy of *Picea crassifolia* and *Sabina przewalskii* for the combination eight

类别	用户精度/%	生产精度/%
青海云杉	86.03	89.79
祁连圆柏	86.61	90.12

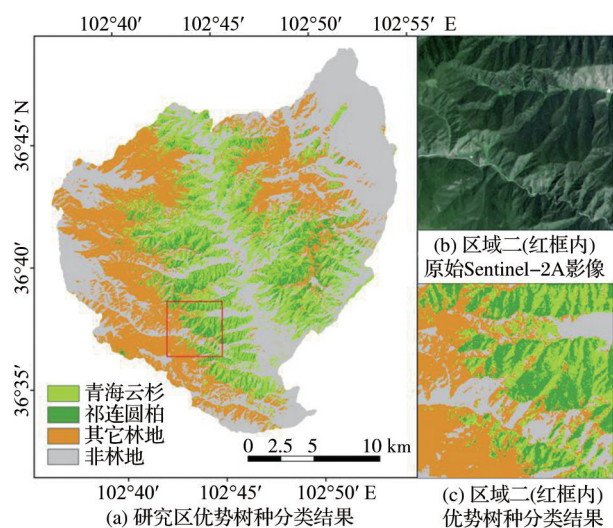


图6 研究区优势树种分类结果及细节对比

Fig.6 The classification result in study area and details comparison

模型建立时还能计算特征重要性,即不同特征变量对最终分类结果的贡献。对分类精度最高的组合8中22个特征变量做归一化特征重要性评分(图7)。

山地森林的森林类型分布与地形因素有很大关系,图7也表明海拔和坡向对保护区优势树种分类的重要性。山区崎岖不平的地表起伏导致影像部分像元的反射率受到影响,阴影问题尤其在春季或冬季的遥感影像中比较严重,选取夏季时间的光学影像,并在分类过程通过加入DEM、坡向等辅助信息,解决阴影对分类结果的影响,避免地形效应^[21-22]。青海云杉和祁连圆柏分类的最高总体精度来自于将VV和VH后向散射系数与Sentinel-2A、Landsat-8和DEM结合,添加VV和VH后向散射系数有助于总体精度小幅度提升,这与Liu等^[23]得到的结果一致。

5 讨论和结论

本文利用Sentinel-2A、Sentinel-1A和Landsat-8共3种多源遥感影像及DEM数据,基于随机森林分类方法,对连城国家级自然保护区内青海云杉和祁连圆柏进行分类,得到以下结论:

(1)与单一影像相比,添加多源遥感影像和DEM数据不论是林地提取还是优势种分类都能明显提高总体精度。Sentinel-2A较Landsat-8分辨率高,波段丰富,在提取林地时添加DEM得到最高精度98.43%,对青海云杉和祁连圆柏分类时,结合地形特征精度有明显提高,基于之上再添加VV和VH后向散射系数精度达到92.85%。

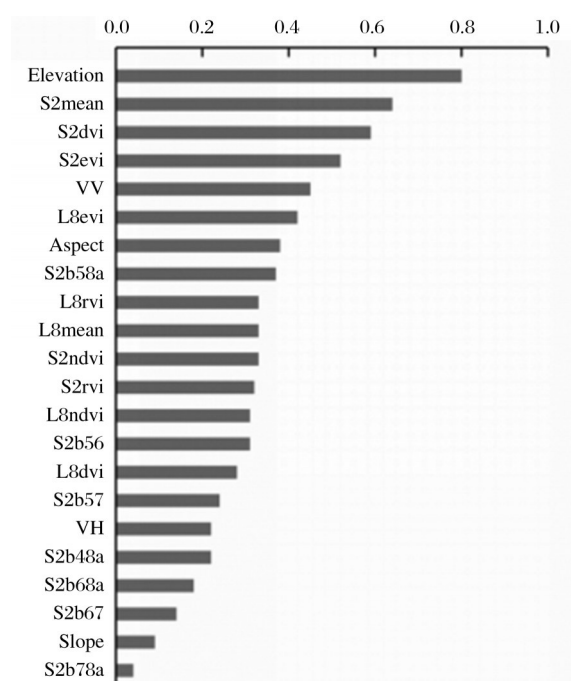


图7 特征变量重要性排序

Fig.7 Importance ranking of feature variables

(2)地形特征中对优势种分类贡献较大的特征变量为海拔和坡向,贡献最小的是坡度,这与青海云杉和祁连圆柏的生长特点有关,青海云杉多分布在山谷、阴坡,祁连圆柏基本分布在阳坡。Sentinel-2A的光谱特征S2mean、植被指数DVI、EVI也具有较高的重要性,而由3个植被红边波段计算的植被指数和Landsat-8的植被指数评分接近。虽然结合Sentinel-1A分类总体精度仅有小幅度提高,但VV极化方式的后向散射系数对分类结果仍有较大贡献。

(3)利用多源遥感影像及数字高程数据,提取光谱特征、植被指数、纹理特征和地形特征,设置不同组合方案,对森林类型进行分层提取,将光学影像的光谱信息和SAR卫星数据的微波信息优势互补,获取比单一数据源更好的分类效果。本文方法在森林类型识别方面有显著优势,在环境复杂的山地森林资源调查、优势树种组识别等方面具有较大潜力。

然而本研究在诸多方面还存在不足,在后续研究工作中需完善并进行深刻讨论:

(1)本文仅基于遥感数据多源化,并未考虑多时相,而植被季相节律信息能通过多时相遥感数据获得,因此后续工作加入多时相的Landsat 8或Sentinel-2A影像,有助于分类精度的再度提高。

(2)随机森林方法可估算特征变量重要性,因

此能根据特征重要性排序优选特征变量,剔除贡献最低的特征变量后再次分类估算新的特征重要性排序,重复步骤进行多次试验以寻求最高精度的分类结果和特征组合。

(3)本文提出的研究方法在连城保护区取得较好的分类结果,后续研究中可考虑扩展到大区域的祁连山保护区,提取较大范围青海云杉和祁连圆柏的空间分布范围,尝试大面积山地森林类型制图。

参考文献(References):

- [1] Wu Y S, Zhang X L. Object-based Tree Species Classification Using Airborne Hyperspectral Images and LiDAR Data [J]. *Forests*, 2020, 11, 32. doi: 10.3390/f1101003.
- [2] Krzystek P, Serebryanyk A, Schnörr C, *et al.* Large-scale Mapping of Tree Species and Dead Trees in Šumava National Park and Bavarian Forest National Park Using LiDAR and Multispectral Imagery [J]. *Remote Sensing*, 2020, 12, 661. doi: 10.3390/rs12040661.
- [3] 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, 1070. doi: 10.3390/rs12071070.
- [4] Ren Chong, Ju Hongbo, Zhang Huaiqing, *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.]
- [5] Li Mengying, Xing Yanqiu, Liu Meishuang, *et al.* Identification of Forest Type with Landsat-8 Image based on SVM [J]. *Journal of Central South University of Forestry & Technology*, 2017, 37(4): 52-58. [李梦颖, 邢艳秋, 刘美爽, 等. 基于支持向量机的 Landsat-8 影像森林类型识别研究[J]. *中南林业科技大学学报*, 2017, 37(4): 52-58.]
- [6] Tian Tian, Fan Wenyi, Lu Wei, *et al.* An Object-based Information Extraction Technology for Dominant Tree Species Group Types [J]. *Chinese Journal of Applied Ecology*, 2015(6): 1665-1672. [田甜, 范文义, 卢伟, 等. 面向对象的优势树种类型信息提取技术[J]. *应用生态学报*, 2015(6): 1665-1672.]
- [7] Markus I, Francesco V, Clement A. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe [J]. *Remote Sensing*, 2016, 8(3): 166. doi: 10.3390/rs8030166.
- [8] Zhu X L, Liu D S. Accurate Mapping of Forest Types Using Dense Seasonal Landsat Time-series [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 96: 1-11.
- [9] Xie Z L, Chen Y L, Lu D S, *et al.* Classification of Land Cover, Forest, and Tree Species Classes with ZiYuan-3 Multispectral and Stereo Data [J]. *Remote Sensing*, 2019, 11(2): 164. doi: 10.3390/rs11020164.
- [10] Watanachaturaporn P, Arora M K, Varshney P K. Multisource Classification Using Support Vector Machines [J]. *Photogrammetric Engineering & Remote Sensing*, 2015, 74(2): 239-246.
- [11] Forster M, Kleinschmit B. Significance Analysis of Different Types of Ancillary Geodata Utilized in a Multisource Classification Process for Forest Identification in Germany [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 52(5): 3453-3463.
- [12] Xu X D, Li W, Ran Q, *et al.* Multisource Remote Sensing Data Classification based on Convolutional Neural Network [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2017, (99): 1-13.
- [13] Li J Y, Zhang H Y, Guo M, *et al.* Urban Classification by the Fusion of Thermal Infrared Hyperspectral and Visible Data [J]. *Photogrammetric Engineering & Remote Sensing*, 2015, 81(12): 901-911.
- [14] Lei Guangbin, Li Ainong, Tan Jianbo, *et al.* Forest Types Mapping in Mountainous Area Using Multi-source and Multi-temporal Satellite Images and Decision Tree Models [J]. *Remote Sensing Technology and Application*, 2016, 31(1): 31-41. [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.]
- [15] Liu Gang, Zhang Yu, Zang Zhuo, *et al.* Extraction and Comparative Analysis of Forest Information of Multi-source Remote Sensing Data [J]. *Journal of Central South University of Forestry & Technology*, 2012, 32(10): 158-161. [刘刚, 张雨, 臧卓, 等. 多源遥感数据森林信息的提取和比较分析[J]. *中南林业科技大学学报*, 2012, 32(10): 158-161.]
- [16] Wen Xiaole, Zhong Ao, Hu Xiujuan. The Classification of Urban Greening Tree Species based on Feature Selection of Random Forest [J]. *Journal of Geo-information Science*, 2018, 20(12): 1777-1786. [温小乐, 钟奥, 胡秀娟. 基于随机森林特征选择的绿化乔木树种分类[J]. *地球信息科学学报*, 2018, 20(12): 1777-1786.]
- [17] Lü Jie, Hao Ningyan, Li Chonggui, *et al.* Identification of Forest Type based on Random Forest and Texture Characteristic [J]. *Remote Sensing Information*, 2017, 32(6): 109-114. [吕杰, 郝宁燕, 李崇贵, 等. 利用随机森林和纹理特征的森林类型识别[J]. *遥感信息*, 2017, 32(6): 109-114.]
- [18] Persson M, Lindberg E, Reese H. Tree Species Classification with Multi-temporal Sentinel-2 Data [J]. *Remote Sensing*, 2018, 10, 1794. doi: 10.3390/rs10111794.
- [19] Li Ruonan, Ou Guanglong, Dai Qinling, *et al.* Forest Types Classification in Shangri La, Northwest Yunnan with Landsat Time series Data based on Google Earth Engine [J]. *Journal of Southwest Forestry University (Natural Sciences)*, 2020, 40(5): 1-12. [李若楠, 欧光龙, 代沁玲, 等. 基于 GEE 和 Landsat 时间序列数据的香格里拉森林类型分类研究[J]. *西南林业大学学报(自然科学版)*, 2020, 40(5): 1-12.]
- [20] Liu Sihan, Shang Xiaming, Ma Ting. The Classification of Forest Types. based on Feature Selection of Random Forest

- [J]. Beijing Surveying and Mapping, 2019, 33(12): 1518-1522.[刘思涵,尚夏明,马婷. 基于随机森林特征选择的森林类型分类[J].北京测绘,2019,33(12):1518-1522.]
- [21] Bie Qiang, He Lei, Zhao Chuanyan, *et al.* Study on Vegetation Information Extraction based on Object-oriented Image Analysis [J]. Remote Sensing Technology and Application, 2014, 29(1): 164-171.[别强,何磊,赵传燕. 基于影像融合和面向对象技术的植被信息提取研究[J]. 遥感技术与应用, 2014, 29(1): 164-171.]
- [22] Wang Xiaoxue, Shen Huitao, Lin Tianmiao, *et al.* Improving the Forest Type Mapping Accuracy in Semiarid Mountainous Areas based on TM Image-Take the West Mountain of Beijing as an Example[J]. Journal of Natural Resources, 2017, 32(7): 1217-1228.[王晓学,沈会涛,林田苗,等. 利用多信息源提高半干旱地区TM影像的森林类型制图精度:以北京西部山区为例[J]. 自然资源学报, 2017, 32(7): 1217-1228.]
- [23] Liu Y A, Gong W S, Hu X Y, *et al.* Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data [J]. Remote Sensing, 2018, 10(6): 946. doi:10.3390/rs10060946.

Classification of *Picea crassifolia* and *Sabina przewalskii* based on Multi-source Remote Sensing Images

Li Meng, Nian Yanyun, Bian Rui, Bai Yanping, Ma Jinhui

(College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China)

Abstract: *Picea crassifolia* and *Sabina przewalskii* are the dominant species in Liancheng Nature Reserve. Extracting the spatial distribution of two types of trees is of great significance for the management and monitoring of forest resources in the reserve. Based on the method of random forest, 22 feature variables in eight combinations from Sentinel-2A (S2), Sentinel-1A (S1), Landsat-8 (L8) three remote sensing images and digital elevation model of SRTM DEM to classify *Picea crassifolia* and *Sabina przewalskii* in Liancheng Nature Reserve of Gansu Province. The results demonstrated that the accuracy of integrating VV and VH backscattering information of sentinels-1A (S1) was the highest, reaching 92.85%, which is 11.64% higher than that of single image Landsat-8. Experiments showed that combining different bands of multi-source remote sensing images is an effective means to improve the classification accuracy of forest types, which is beneficial to forest resource survey and vegetation information extraction in complex mountainous areas.

Key words: Sentinel-2A; Feature variables; Random forest; Information extraction; Liancheng National Nature Reserve