

盛艳芳, 买买提·沙吾提, 何旭刚, 等. 基于 Google earth engine 渭-库绿洲果园遥感提取[J]. 江苏农业学报, 2024, 40(1): 103-111.
doi: 10.3969/j.issn.1000-4440.2024.01.011

基于 Google earth engine 渭-库绿洲果园遥感提取

盛艳芳^{1,2,3}, 买买提·沙吾提^{1,2,3}, 何旭刚^{1,2,3}, 李荣鹏^{1,2,3}

(1.新疆大学地理与遥感科学学院, 新疆 乌鲁木齐 830017; 2.新疆大学新疆绿洲生态重点实验室, 新疆 乌鲁木齐 830017;
3.新疆大学智慧城市与环境建模自治区普通高校重点实验室, 新疆 乌鲁木齐 830017)

摘要: 针对干旱区果园大面积遥感提取困难、识别精度低等问题, 本研究基于 GEE (Google earth engine) 平台, 综合应用 Sentinel-1/Sentinel-2 影像构建特征集。通过对比原始特征组合、Jeffries-Matusita (J-M) 距离、属性重要度 3 种优化方式, 结合随机森林 (Random forest, RF) 分类方法, 对比得到最佳优化方式, 探索果园最优分类特征集。结果表明: 识别效果最好的方案为 G17_{JM}, 总体精度为 91.25%, kappa 系数为 0.89, 面积精度为 82.55%。最优特征集为 B8_{asm}、B8_{ent}、B8_{idm}、NDVI_{re3}、B6、B7、a、e、b、EVI、B11、B8A、B8、VV。使用 J-M 距离进行特征集优化, 有效降低数据量、提高计算效率, 更有利于精确遥感识别果园种植面积。表明 GEE 快速、准确获取果园种植面积的可行性, 为获取果园动态变化提供强有力的基础。

关键词: Google earth engine (GEE); 特征优化; J-M 距离; 特征集

中图分类号: S127 **文献标识码:** A **文章编号:** 1000-4440(2024)01-0103-09

Remote sensing extraction of orchard in the casis of weigan and kuqa rivers based on Google earth engine

SHENG Yan-fang^{1,2,3}, SAWUT Mamat^{1,2,3}, HE Xu-gang^{1,2,3}, LI Rong-peng^{1,2,3}

(1. College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi 830017, China; 2. Xinjiang Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi 830017, China; 3. Key Laboratory of Smart City and Environment Modelling of Higher Education Institute, Xinjiang University, Urumqi 830017, China)

Abstract: Aiming at the problems of difficult extraction and low recognition accuracy of orchards in arid areas, based on the Google earth engine (GEE) platform, this study comprehensively applied Sentinel-1/Sentinel-2 images to construct feature sets. By comparing the three optimization methods of original feature combination, Jeffries-Matusita (J-M) distance and attribute importance, combined with random forest (RF) classification method, the best optimization method was obtained, and the optimal classification feature set of orchard was explored. The results showed that the best recognition scheme was G17_{JM}, the overall accuracy was 91.25%, kappa coefficient was 0.89, and area accuracy was 82.55%. The optimal feature set was B8_{asm}, B8_{ent}, B8_{idm}, NDVI_{re3}, B6, B7, a, e, b, EVI, B11, B8A, B8, VV. Using J-M distance to optimize the feature set can effectively reduce the amount of data and improve the computational efficiency, which was more conducive to the accurate identification of orchard planting area. It shows that GEE is feasible to obtain orchard

planting area quickly and accurately, and provides a strong basis for obtaining orchard dynamic changes.

Key words: Google earth engine (GEE); feature optimization; J-M distance; feature set

收稿日期: 2023-02-05

基金项目: 新疆自然科学基金 (自然科学基金) 联合基金项目 (2021D01C055)

作者简介: 盛艳芳 (1993-), 女, 新疆库尔勒人, 硕士研究生, 主要从事农业遥感分类等方面研究。(E-mail) syf2258164@fox-mail.com

通讯作者: 买买提·沙吾提, (E-mail) korxat@xju.edu.cn

新疆是世界六大果品生产带之一和久负盛名的“瓜果之乡”^[1]。自 2018 年开始以“新疆特色林果提

质增效,助力脱贫攻坚”举措^[2]为基础,全力推进林业扶贫工作,南疆林果业规模快速扩张,林果面积每年增加百万亩,形成环塔里木盆地林果主产区等林果基地,在推进新疆脱贫攻坚方面具有重要作用。然而,南疆林果业种植面积大、地块较为破碎,果粮间作种植模式普遍存在,传统的果园种植面积监测方式已经难以满足当前林果业快速发展的实际需求。

近年来,随着遥感(多光谱、中高分辨率影像)的广泛使用,以及机器学习算法、谷歌地球引擎(Google earth engine, GEE)云计算平台不断发展,遥感技术在林果业信息提取、长势监测和估产等方面发挥了重要作用^[3-6]。利用 Sentinel^[7]、HJ^[8]、GF^[9]系列等卫星影像,采用随机森林(RF)^[10]、支持向量机(SVM)^[11]、BP神经网络(BP)^[12]等分类技术对小麦、玉米等常见的农作物进行种植面积信息提取^[13]。RF分类在调查果园分布的应用和研究中广泛使用^[14-17]。另外,大部分研究主要集中在江西^[18]、浙江^[19]、福建^[20]等丘陵地区,并且研究对象也仅局限于柑橘等芸香科果树。目前对于西北干旱

区果树,尤其是新疆地区果园的遥感监测缺乏深入研究。

基于现状,还需要加强对果树遥感分类的研究,不断增加遥感影像、改进面积与分布信息的提取方法以适应技术发展要求,实现快速提取果园面积空间分布状况的目的。本研究通过中高分辨率卫星影像,充分挖掘多源卫星遥感特征量,发展基于 GEE 平台和机器学习的果园智能识别技术,突破高精度果园空间分布遥感制图的技术瓶颈,为果园种植面积空间分布调查和定期动态更新提供技术支持。

1 材料与方法

1.1 研究区概况

渭干河-库车河三角洲绿洲(渭-库绿洲)是目前新疆重要的林果产业基地,位于新疆塔里木盆地(图1),是典型的冲积扇绿洲,属暖温带大陆性干旱气候,平均气温 10.0 °C,日照 2 845~2 977 h/a,光热资源充裕,年降水量 59 mm。相对充足的水分、热量,形成得天独厚的光、热等自然资源,有利于果树生长。

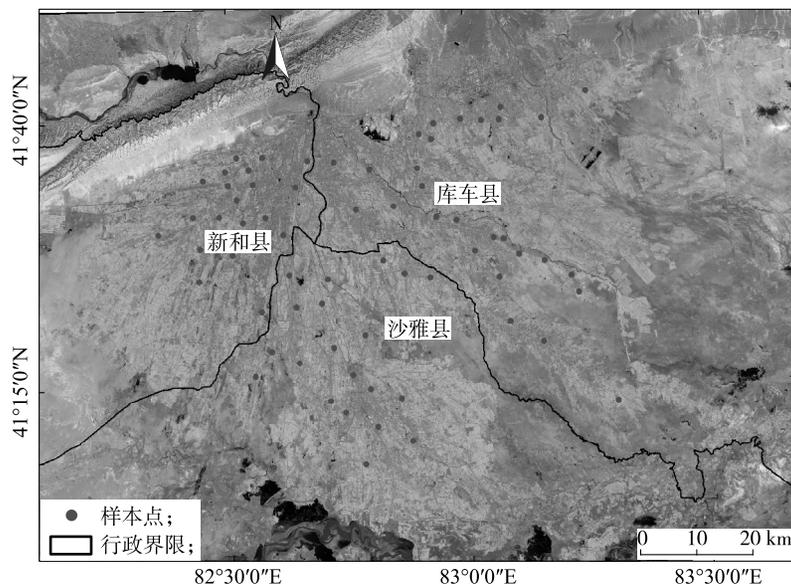


图1 研究区示意图

Fig.1 Map of study area

1.2 研究数据

遥感数据源于 GEE 平台,包括 Sentinel-2 多光谱影像和 Sentinel-1 雷达影像(表 1)。预处理包括:轨道参数标定、热噪声去除、辐射定标、地形校正、后向散射系数提取等,通过 GEE 内嵌算法统一坐标系确保不同数据源之间的几何配准精度,空间分辨率

统一重采样至 10 m。使用平台 JavaScript 语言构建模型进行果园信息提取识别。

野外数据:分别在 2021 年 5 月 26 日至 6 月 3 日、2022 年 6 月 28 日至 7 月 7 日进行 2 次野外实地考察,结合研究区域地物类型以及相关国家标准(GB/T 21010-2017)确定研究区地物为:果园、建

设用地、林地、草地、水域、耕地、其他地物,共获得典型地物样点 920 个。此外,使用 Google earth 影像结合野外获得的典型地物样本点进行目视解译来扩充样本数量,使样本点均匀分布在整个绿洲内部,最终样本总数为1 139个,训练集和验证集比例为7:3。

表 1 本研究使用的遥感数据

Table 1 Remote sensing data used in this study

卫星	波段数量	空间分辨率 (m)	影像数量	成像时间
Sentinel-2	13	10	343	2019-01-07、...、2021-12-22
Sentinel-1	2	10	12	2020-01-12、...、2020-12-02

1.3 研究方法

本研究基于 GEE 平台,使用 Sentinel-1 和 Sentinel-2 数据,计算得到 6 类特征,使用 3 种特征优化方法结合随机森林分类(Random forest, RF),确定提取果园信息的最优特征集(图 2)。

1.3.1 分类特征参数和特征优化方法 通过计算得到光谱特征、植被指数特征、红边特征、纹理特征、极化特征和物候特征构建分类特征集^[21],共包含 42 个子特征(表 2)。

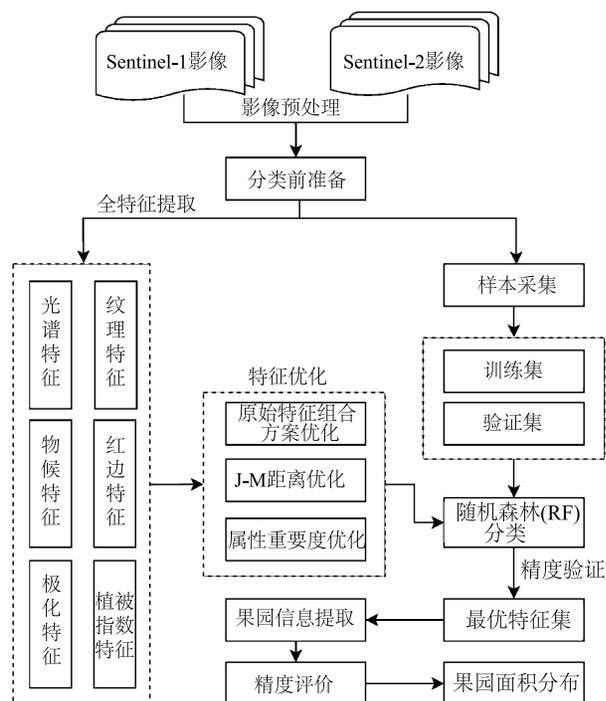


图 2 技术路线图

Fig.2 Technology roadmap

表 2 分类特征参数

Table 2 Classification characteristic parameters

特征集	特征参数	说明或计算公式
光谱特征(B)	光谱特征参数	B1、B2、B3、B4、B8、B8a、B9、B10、B11、B12
植被指数特征(Z)	归一化差异植被指数(NDVI)	$(B8a-B4)/(B8a+B4)$
	增强型植被指数(EVI)	$\frac{2.5 \times (B8a-B4)}{B8a+6 \times B4-7.5 \times B2+1}$
	归一化水体指数(NDWI)	$(B3-B8a)/(B3+B8a)$
红边特征(H)	归一化差异红边指数 1(NDVI _{re1})	$(B8a-B5)/(B8a+B5)$
	归一化差异红边指数 2(NDVI _{re2})	$(B8a-B6)/(B8a+B6)$
	归一化差异红边指数 3(NDVI _{re3})	$(B8a-B7)/(B8a+B7)$
	归一化差异红边边缘指数 1(ND _{re1})	$(B6-B5)/(B6+B5)$
	归一化差异红边边缘指数 2(ND _{re2})	$(B7-B5)/(B7+B5)$
	红边波段(B _{re})	B5、B6、B7
纹理特征(T)	B8 _{asm}	角二阶矩
	B8 _{contrast}	对比度
	B8 _{corr}	相关性
	B8 _{var}	方差
	B8 _{idm}	逆差距
	B8 _{diss}	差异性
	B8 _{sent}	和熵
	B8 _{ent}	熵
极化特征(L)	VV	入射波为垂直偏振时的垂直偏振后向散射
	VH	入射波为垂直偏振时的水平偏振后向散射
物候特征(W)	a	生长季开始
	b	生长季结束
	c	生长季长度
	d	基值
	e	生长季中期
	f	生长季拟合函数最大值
	g	生长季幅度
	h	生长季左导数
	i	生长季右导数
	j	生长季小积分
	k	生长季大积分

1.3.2 特征集优化

1.3.2.1 基于 J-M 距离的特征集优化 J-M 距离衡量不同地物类型在波段间分离能力,对各地物类型的可分性进行定量分析,并据此进行特征集优

化^[22]。表达式为:

$$JM = 2(1 - e^{-B})$$

式中 B 表示某一特征 的巴氏距离 (Bhattacharyya distance)。不同种类样本的巴氏距离同时 样本满足正态分布的时为

$$B = \frac{1}{8}(e_1 - e_2)^2 \frac{2}{\delta_1^2 + \delta_2^2} + \frac{1}{2} \ln \left(\frac{\delta_1^2 + \delta_2^2}{2\delta_1^2 \cdot \delta_2^2} \right)$$

式中 e_k 表示某类特征的均值, δ_k^2 表示某类特征的方差。

1.3.2.2 基于随机森林属性重要度的特征集优化 随机森林分类器 (RF) 是一种非参数分类器, 由一组决策树组成, 具有便捷的特征选择算法^[23]。利用 RF 中的属性重要度, 进行有效数据降维, 根据得分由高至低对特征集进行优化。

1.3.2.3 分类方案制定和精度验证 本研究利用 GEE 平台中最优棵树选择, 最终确定待生成的决策树 (Ntree) 数目值为 100。为了说明有利于果园信息提取的最佳特征, 设计了 23 种试验方案 (表 3)。

表 3 分类方案

Table 3 Classification scheme

方案	特征集/说明
G1	B、Z
G2	B、H
G3	B、L
G4	B、T
G5	B、W
G6	B、Z、L
G7	B、Z、T
G8	B、Z、W
G9	B、L、T
G10	B、L、W
G11	B、T、W
G12	B、T、L、W
G13	B、H、T、W
G14	B、Z、L、T
G15	B、Z、L、W
G16	B、Z、T、W
G17	B、Z、H、L、W、T
G4 _{JM} 、G12 _{JM} 、G17 _{JM}	J-M 距离优化 G4、G12、G17
G4 _{VIP} 、G12 _{VIP} 、G17 _{VIP}	属性重要度优化 G4、G12、G17

B、Z、T、H、L、W 见表 2。

1.3.3 精度验证 研究表明, 总体精度 (OA) 与 Kappa 系数在遥感分类精度验证中是最基本、最能

衡量分类结果的指标^[24]。其中 OA 是指被正确分类的像元之和占有地类像元总数的百分比, Kappa 系数表明分类结果与验证样本数据之间的一致性。因此本研究使用 OA、Kappa 系数作为总体分类精度评价体系, 使用统计年鉴数据对提取的果园种植面积信息进行验证。

2 结果与分析

2.1 果园物候特征分析

为了确定果园信息提取的最佳时间窗口, 使用 Sentinel-2 数据, 综合研究区域典型地物样本点建立归一化植被指数 (NDVI) 时间序列曲线, 以及使用 Timesat3.3 软件提取植被物候特征。

NDVI 时间序列曲线, 反映了植被与非植被在 2020 年 NDVI 变化特征^[25]。如图 3 所示, 非植被 (建设用地、裸地、水域) 的 NDVI 值全年较低。3 月植被 (果园、林地、耕地、草地) NDVI 最小 (NDVI < 0.1); 3-6 月植被 NDVI 开始快速上升, 其中果园上升速度最快, NDVI 最高, 尤其在 5 月远高于其他地物 (NDVI > 0.35); 6-8 月, 植被 NDVI 逐渐增长至峰值, 即生长旺盛期; 8 月裸地、建设用地、草地、果园、林地、耕地 NDVI 均开始下降, 果园与林地 NDVI 相近。

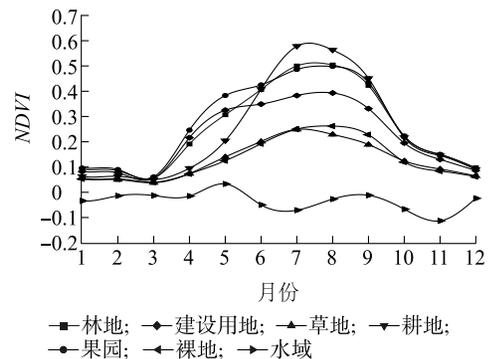
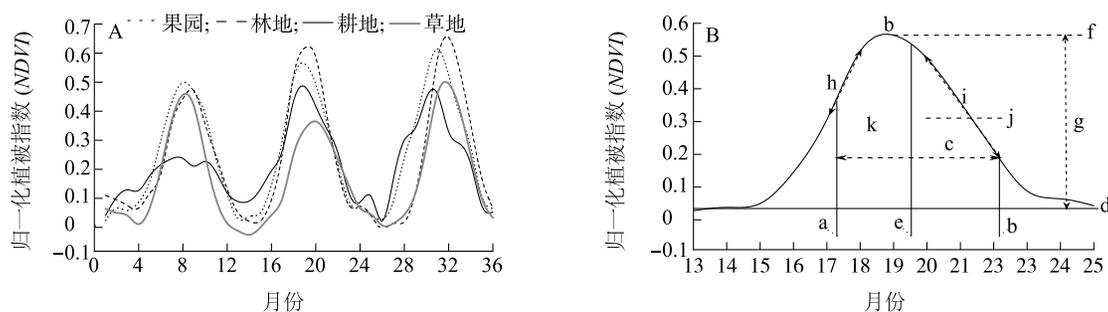


图 3 2020 年归一化植被指数 (NDVI) 时间序列曲线

Fig.3 Time series curve of normalized difference vegetation index (NDVI) in 2020

植被物候信息如图 4A 所示, 5 月果园 NDVI 高于所有其他植被, 结合图 4B 可知果园生长季在 5 月 (a 点) 开始, 此时果园在所有植被中 NDVI 最大 (NDVI > 0.4); 而后所有植被的 NDVI 继续增加至生长季中期 (6-7 月), 7 月中旬所有植被的 NDVI 达到最大值, 即各植被均到达生长旺盛期 (NDVI > 0.4); 8 月开始所有植被的 NDVI 逐渐下降, 10 月生长季结束, NDVI 迅速下降。



A 图为 4 种主要植被物候曲线图;B 图为果园物候信息点图。物候特征 a~k 见表 2。

图 4 4 种主要植被物候曲线及果园物候信息点

Fig.4 Four main vegetation phenological curves and orchard phenological information points

为了探讨雷达数据在果园信息提取中的作用,基于 2020 年全年 12 期的 Sentinel-1 雷达遥感影像,由 VV、VH 极化数据算出典型地物的后向散射系数,得到典型地物后向散射特征时间序列(图 5)。类型不同的地物,散射特征随着生长发育期的变化表现出不同的散射机制^[26]。

从时间尺度,以及不同极化方式上来看,VV 后向散射系数变化特征相比于 VH 显著。果树的萌芽期和展叶期在 4-5 月,在 VV 极化中果园与耕地区分不明显,趋势几乎一致,在 VH 极化中 5 月果园略高于

耕地。本研究果园中的果树与林地中的树木存在一定的相似性,极化数据在不同类型作物的识别中具有较好的识别效果^[27];林地在 VV 极化方式下,1-3 月与果园的后向散射系数趋势一致,在 4 月开始有所区分,林地后向散射系数高于果园,峰值在 6 月达到 -6.99 dB,在 VH 极化中,至 8 月前都有较好的分离性,因此提取果园信息的最佳时间点是在 5-7 月。

最后,综合 Sentinel-1 极化影像后向散射特征和 Sentinel-2 多光谱影像计算得到的 NDVI 时间序列与物候特征,可以确定 5 月为果园信息提取的最佳时间窗口。

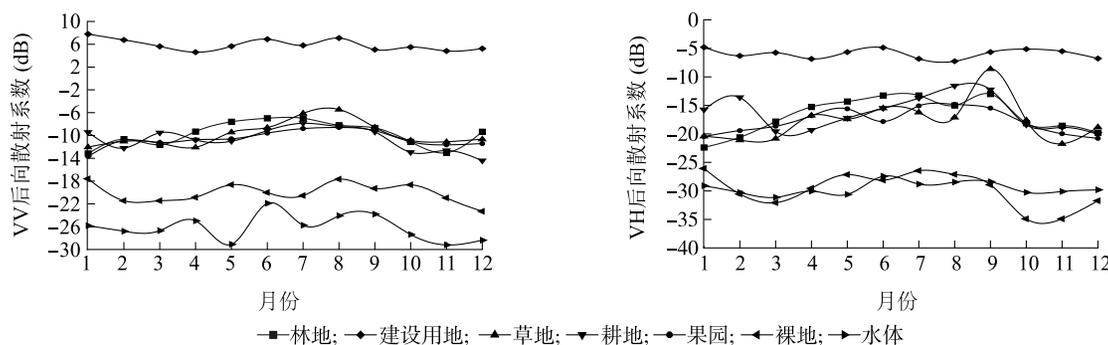


图 5 2020 年 VV 与 VH 极化后向散射系数时间序列曲线

Fig.5 Time series curves of VV and VH polarized backscattering coefficients in 2020

2.2 特征集优化分析

2.2.1 分类方案特征集优选 所有特征排列组合形成不同的分类方案,并通过筛选得到 17 个特征方案(表 3),使用 RF 分类(图 6)。在原始特征组合而成的分类方案中,所有方案总体精度均达到了 80% 以上,总体精度排列前三的方案分别是 G12、G4、G17。

2.2.2 基于 J-M 距离与属性重要度的特征集优化 获取 J-M 距离和属性重要度对各特征的评分,依据 2 种评价方法,依次加入特征,观察果园的分类总体精度变化(图 7)。2 种方法表现出一些类似的特点,随着所用

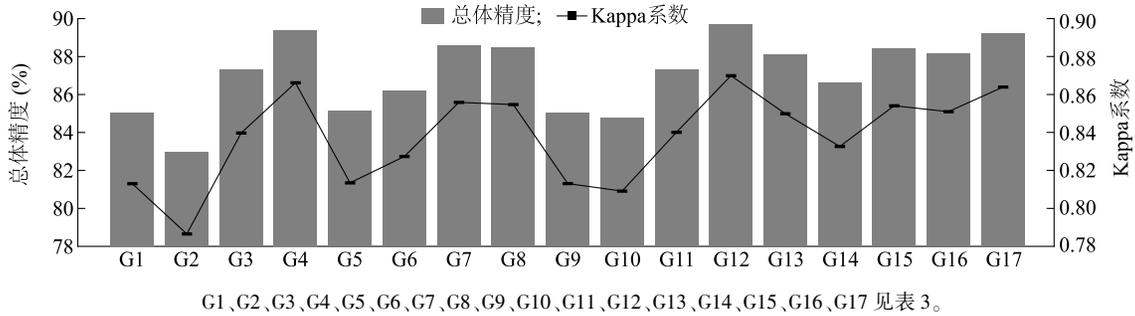
特征数逐渐增加,总体精度表现为快速增长,当特征数达到 10 即 B6 处时,总体精度开始趋于稳定,特征数达到 40 即 B8_contrast 时能达到最优精度约 89%。

根据图 7,按照 6 个特征类别进行分类统计(表 4)。通过对比,在两种优化的方法中,纹理特征的可分离性,对分类贡献度均较高,纹理作为物体表面的一种基本属性广泛存在自然界中,是描述和识别物体的一种极为重要的特征^[28],果园因其均匀、粗糙等纹理特征明显不同于其他地类,尤其是与林地间的差别较大。

2 种方法中排名靠前的物候特征是 a、b、e,即生长

季开始、生长季结束、生长季中期,它们可以直接反映植被的生长信息。所以纹理特征与物候特征相比其他

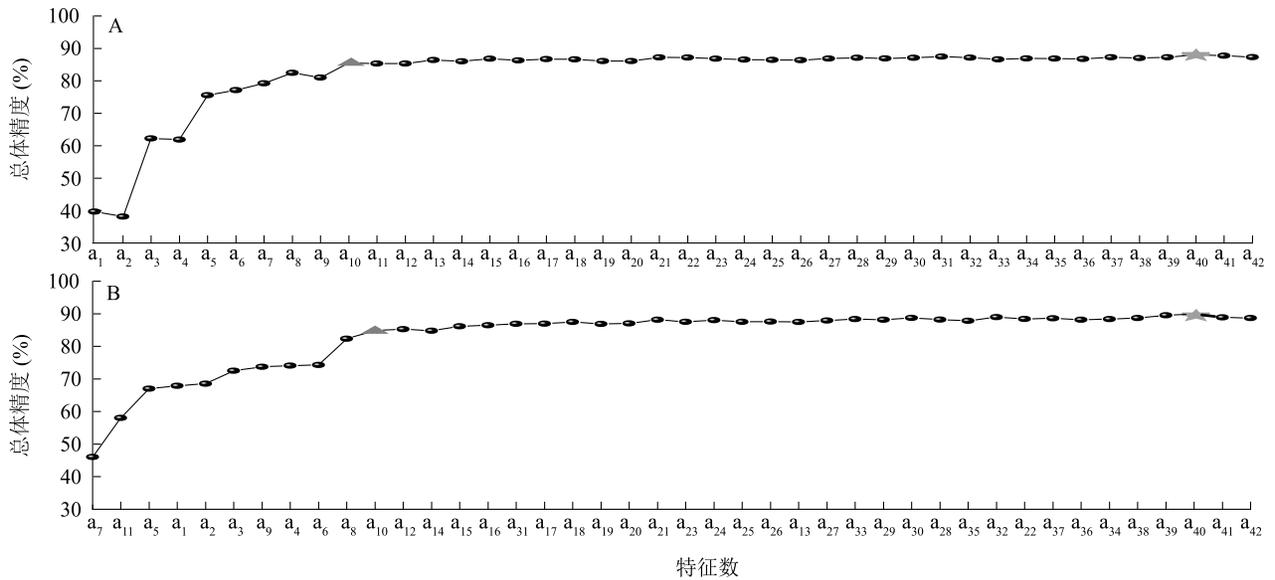
特征更能凸显不同植被间的差异,更好突出果园与其他植被之间的差异^[29]。



G1、G2、G3、G4、G5、G6、G7、G8、G9、G10、G11、G12、G13、G14、G15、G16、G17 见表 3。

图 6 分类方案的总体精度

Fig.6 The overall accuracy of classification scheme



A: J-M 距离优化方法; B: 属性重要度优化方法。a₁: B8_asm; a₂: B8_ent; a₃: NDVI_{re3}; a₄: B8_idm; a₅: a; a₆: B8_sent; a₇: e; a₈: EVI; a₉: d; a₁₀: B6; a₁₁: b; a₁₂: B11; a₁₃: f; a₁₄: NDWI; a₁₅: VV; a₁₆: B8a; a₁₇: B7; a₁₈: B8; a₁₉: NDVI_{re1}; a₂₀: B12; a₂₁: VH; a₂₂: NDVI; a₂₃: ND_{re2}; a₂₄: ND_{re1}; a₂₅: B5; a₂₆: B3; a₂₇: B4; a₂₈: B1; a₂₉: B9; a₃₀: NDVI_{re2}; a₃₁: j; a₃₂: g; a₃₃: B2; a₃₄: i; a₃₅: k; a₃₆: h; a₃₇: c; a₃₈: B8_var; a₃₉: B8_diss; a₄₀: B8_contrast; a₄₁: B10; a₄₂: B8_corr。a、b、c、d、e、f、g、h、i、j、k、NDVI_{re1}、NDVI_{re2}、NDVI_{re3}、ND_{re1}、ND_{re2}、B1、B2、B3、B4、B5、B6、B7、B8、B8a、B9、B10、B11、B12、B8_asm、B8_contrast、B8_corr、B8_var、B8_idm、B8_diss、B8_sent、B8_ent、NDVI、EVI、NDWI、VV、VH 见表 2。

图 7 总体精度变化图

Fig.7 Overall accuracy variation map

表 4 特征优化表

Table 4 Feature optimization table

优化方法	纹理特征	红边特征	物候特征	植被指数	光谱特征	极化特征
J-M 距离	B8_asm	NDVI _{re3}	a	EVI	B11	VV
	B8_ent	B6	e	/	B8a	/
	B8_idm	B7	d	/	B8	/
属性重要度	B8_asm	NDVI _{re3}	e	EVI	B11	VV
	B8_ent	B6	b	/	B8a	/
	B8_idm	B7	a	/	B8	/

B8_asm、B8_ent、B8_idm、NDVI_{re3}、EVI、VV、B6、B7、B8、B8a、B11、a、b、e、d 见表 2。

2.3 最佳分类方案精度评价

按照优化后的特征集,对原始特征组合中精度较高的方案 G4、G12、G17 依次优化。如表 5 所示,其中精度最高的方案 G17_{JM},总体精度比优化前的方案高 2.05 个百分点,其次方案 G12_{JM}的总体精度比 G12 高 1.16 个百分点;方案 G4、G4_{JM}、G4_{VIP}总体精度均未达到 90%,因此参与的特征数较少时,识别效果较差。

总体而言,使用 J-M 距离优化的方案总体精度均

较高,相比于原始特征组合方案有1~2个百分点的提升,比使用属性重要度优化的高1~3个百分点,优化效果较好。

表 5 分类精度对比

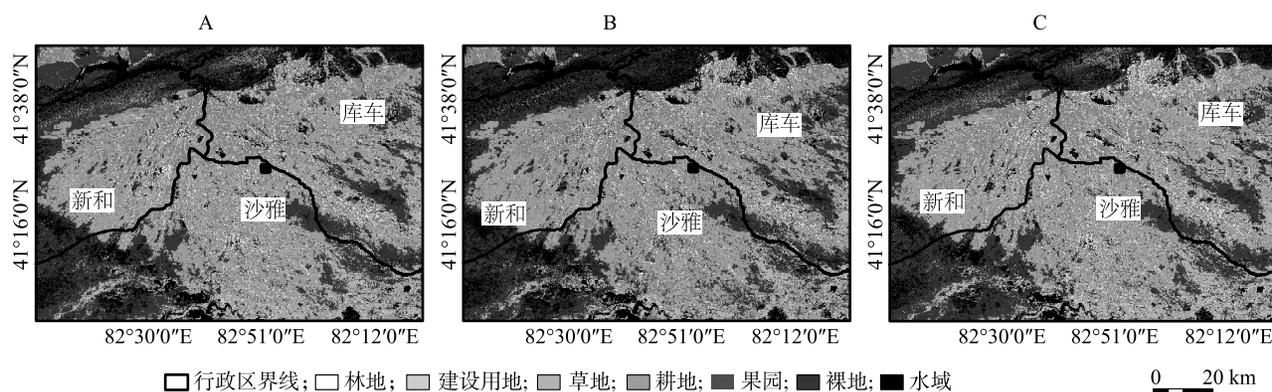
Table 5 Comparison of classification accuracy

方案	总体精度(%)	Kappa 系数
G4	82.95	0.79
G12	89.67	0.87
G17	89.20	0.86
G4 _{JM}	78.21	0.73
G12 _{JM}	90.83	0.88
G17 _{JM}	91.25	0.89
G4 _{VIP}	81.37	0.77
G12 _{VIP}	89.13	0.86
G17 _{VIP}	90.21	0.88

G4、G12、G17、G4_{JM}、G12_{JM}、G17_{JM}、G4_{VIP}、G12_{VIP}、G17_{VIP}见表3。

精度最高的3种方案如图8所示,从空间分布上来看,果园主要分布在库车市和新和县;沿河流与农田交错分布,呈片状、由上至下扇形分布在绿洲中上部。

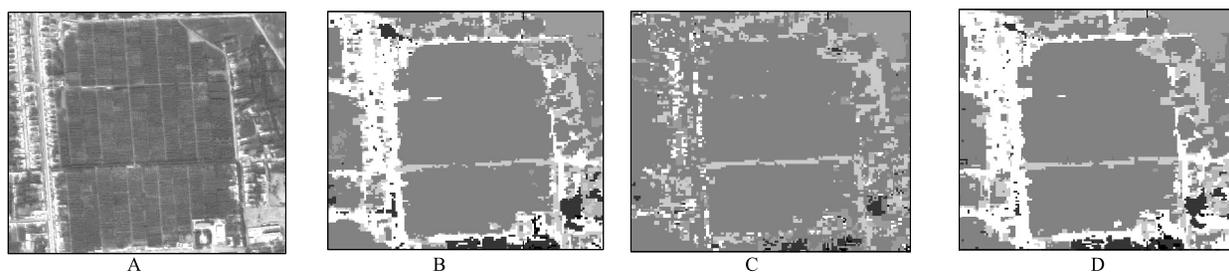
为了比较分类效果,选取经实地考察并包括果园周围常见地类(农田、林地、建设用地等)的局部区域与高分影像(1 m)对比分析(图9)。G12方案与G17_{VIP}方案识别效果相似,可以完整识别果园地块,但部分其他用地被错分为建设用地,林地识别不完整,草地提取范围过大;G17_{JM}方案识别效果更精细,果园地块周围的耕地、草地识别完整,可以将果园与林地准确分开。分类识别的果园种植面积为66 921.62 hm²,统计年鉴^[30]中库车、沙雅、新和果园总面积为81 066.67 hm²,面积精度为82.55%。故J-M距离优化特征的方案G17_{JM}总体精度最高,识别效果最好。



A、B、C 分别为 G12、G17_{JM}、G17_{VIP} 方案提取的果园信息空间分布。G12、G17_{JM}、G17_{VIP}见表3。

图8 提取的果园空间分布图

Fig.8 Spatial distribution map of the extracted orchard information



A 为高分原始影像;B、C、D 分别为 G12、G17_{JM}、G17_{VIP} 方案提取的果园信息空间分布细节图。G12、G17_{JM}、G17_{VIP}见表3。

图9 提取的果园信息空间分布细节图

Fig.9 Spatial distribution detail map of the extracted orchard information

3 讨论

利用 GEE 平台强大的数据运算和存储能力,使用 Sentinel-1/2 遥感数据和 RF 分类算法,在果园信息提取中对特征进行优化的研究相对较少,大多研究集中在农作物监测方面^[31],随着 GEE 平台的发展,逐渐出现不同的研究方向如湖泊面积测量^[32]、沙漠化监测^[33]、黄河大尺度面积制图^[34]等,均取得较好的效果。渭库绿洲作为典型干旱区绿洲,是新疆重要的林果产业基地。种植的人工林区域(果园)与自然林地存在区别,果园地块破碎但边界整齐,纹理规律,经实地调查,以及综合采集多种果园地物,可以发现不同类型果树物候期存在细小差别,茂盛期基本一致,可以利用茂盛期相似的纹理特征、极化特征等,对果园信息进行提取。赵安周等^[35]、徐超等^[36]、张祯祺等^[37]、宁晓刚等^[38]从中尺度上进行的动态监测研究结果,以及本试验为实现大面积果园动态监测所做的研究结果,都证明了果园信息泛化、统一大面积提取的可行性。但是根据以往针对果园的研究中^[14-16],大多使用传统的监督分类方法进行面积信息提取,在今后的研究中可以考虑更多方法的对比研究,如面向对象结合决策树、面向对象结合随机森林等方法;此外, Sentinel-2 数据空间分辨率为 10 m,针对研究区域内相对较小的破碎果园地块无法识别,可以尝试精度更高的遥感影像,如 GF_2 等,得到更高的制图精度。

从研究结果看,对比 3 种优化方法, J-M 距离优化方法可以有效提高识别效果,这与宁晓刚等^[38]的研究结果一致,充分说明使用 J-M 距离是进行特征优化时优于属性重要度优化方法。此外,通过总体精度变化图发现,适用于果园信息提取的最优特征集应最少包含 10 个子特征,当特征数达到 10 时,总体精度开始趋于稳定,分类可以取得较好的效果。

研究结果表明,使用 RF 识别果园时物候、纹理、极化等特征结合 J-M 距离进行特征集优化可以降低数据量、提高计算效率,使用得到的最优特征集进行分类,总体精度与面积精度均高于 80%,效果较好。本研究基本都在 GEE 平台中进行并获得了较好的分类总体精度,证明了使用 GEE 实时获取渭库绿洲果园面积动态变化的可行性,为果园面积动态监测提供强有力的基础。

4 结论

本研究结论如下:(1)本研究共计包括 23 种分类方案,对比方案精度,所有精度均到达 80%以上,其中精度最高的 G17_{JM} 方案,其总体精度为 91.25%, Kappa 系数为 0.89,面积精度为 82.55%;(2)综合使用 NDVI 时间序列曲线、物候曲线和极化时序曲线确定的窗口为 5 月,即 5 月为果园信息提取的最佳时间;(3)综合对比 3 种优化方法, J-M 距离为最佳优化方法,得到的优化方案(G17_{JM})总体精度较高,且效果最好。提取果园信息的最优特征集,具体为: B8_{asm}、B8_{ent}、B8_{idm}、NDVI_{re3}、B6、B7、a、e、b、EVI、B11、B8A、B8、VV。

参考文献:

- [1] 唐治,杨志华,曹占洲,等. 南疆环塔里木盆地林果种植区近 50a 气候变化特征及对果树生长的影响[J]. 农业网络信息, 2014(4): 27-31.
- [2] 李婷. 新疆林果业助力脱贫攻坚的思考研究[J]. 新疆林业, 2019(4): 11-14.
- [3] 桑国庆,唐志光,毛克彪,等. 基于 GEE 云平台与 Sentinel 数据的高分辨率水稻种植范围提取——以湖南省为例[J]. 作物学报, 2022, 48(9): 2409-2420.
- [4] 王涵涌,张泽,康孝岩,等. 基于 Sentinel-2A 的棉花种植面积提取及产量预测[J]. 农业工程学报, 2022, 38(9): 205-214.
- [5] 周珂,柳乐,张俨娜,等. GEE 支持下的河南省冬小麦面积提取及长势监测[J]. 中国农业科学, 2021, 54(11): 2302-2318.
- [6] TIAN H F, HUANG N, NIU Z, et al. Mapping winter crops in China with multi-source satellite imagery and phenology-based algorithm[J]. Remote Sensing, 2019, 11(7): 820-843.
- [7] 张东彦,杨玉莹,黄林生,等. 结合 Sentinel-2 影像和特征优选模型提取大豆种植区[J]. 农业工程学报, 2021, 37(9): 110-119.
- [8] 朱永森,曾永年,张猛. 基于 HJ 卫星数据与面向对象分类的土地利用/覆盖信息提取[J]. 农业工程学报, 2017, 33(14): 258-265.
- [9] 张悦琦,李荣平,穆西哈,等. 基于多时相 GF-6 遥感影像的水稻种植面积提取[J]. 农业工程学报, 2021, 37(17): 189-196.
- [10] 马战林,刘昌华,薛华柱,等. GEE 环境下融合主被动遥感数据的冬小麦识别技术[J]. 农业机械学报, 2021, 52(9): 195-205.
- [11] 熊皓雨,周小成,汪小钦,等. 基于 GEE 云平台的福建省 10 m 分辨率茶园专题空间分布制图[J]. 地球信息科学学报, 2021, 23(7): 1325-1337.
- [12] 孙少杰,吴门新,庄立伟,等. 基于 CNN 卷积神经网络和 BP 神经网络的冬小麦县级产量预测[J]. 农业工程学报, 2022, 38(11): 151-160.

- [13] 冯志敏,刘兆旭,辛海强. 遥感技术在林果种植区调查中的应用研究[J]. 测绘与空间地理信息,2013,36(10):83-87.
- [14] 林娜,王伟,王斌. 基于随机森林和 Landsat8 OLI 影像的脐橙果园种植信息提取[J]. 地理空间信息,2021,19(11):96-100.
- [15] 钟琪,罗津,齐述华. 随机森林分类算法提取柑橘果园的样本数量敏感性分析[J]. 江西科学,2019,37(5):664-669.
- [16] 徐晗泽宇,刘冲,齐述华,等. 基于随机森林算法的赣南柑橘果园遥感信息提取[J]. 江西师范大学学报(自然科学版),2018,42(4):434-440.
- [17] 宋荣杰,宁纪锋,常庆瑞,等. 基于小波纹理和随机森林的猕猴桃果园遥感提取[J]. 农业机械学报,2018,49(4):222-231.
- [18] 徐晗泽宇,刘冲,王军邦,等. Google Earth Engine 平台支持下的赣南柑橘果园遥感提取研究[J]. 地球信息科学学报,2018,20(3):396-404.
- [19] 王斌,何丙辉,林娜,等. 基于随机森林特征选择的茶园遥感提取[J]. 吉林大学学报(工学版),2022,52(7):1719-1732.
- [20] 陈钰嘉. 基于高光谱和雷达数据特征融合的茶园提取研究[D]. 北京:中国地质大学,2021.
- [21] 赵英时. 遥感应用分析原理与方法[M]. 北京:科学出版社,2013.
- [22] 张猛,曾永年,朱永森. 面向对象方法的时间序列 MODIS 数据湿地信息提取——以洞庭湖流域为例[J]. 遥感学报,2017,21(3):479-492.
- [23] 马玥,姜琦刚,孟治国,等. 基于随机森林算法的农耕地土地利用分类研究[J]. 农业机械学报,2016,47(1):279-303.
- [24] 王刚. 基于 Landsat 遥感图像的山地丘陵区果园识别与监测[D]. 昆明:云南师范大学,2021.
- [25] 郭交,尉鹏亮,周正舒,等. 基于时变特征的多时相 PolSAR 农作物分类方法[J]. 农业机械学,2017,48(12):174-182.
- [26] 李宇宸,张军,薛宇飞,等. 基于 Google Earth Engine 的中老缅交界区橡胶林分布遥感提取[J]. 农业工程学报,2020,36(8):174-181.
- [27] 古丽努尔·依沙克,买买提·沙吾提,马春玥. 基于多时相双极化 SAR 数据的作物种植面积提取[J]. 作物学报,2020,46(7):1099-1111.
- [28] 高程程,惠晓威. 基于灰度共生矩阵的纹理特征提取[J]. 计算机系统应用,2010,19(6):195-198.
- [29] 薛朝辉,钱思羽. 融合 Landsat 8 与 Sentinel-2 数据的红树林物候信息提取与分类[J]. 遥感学报,2022,26(6):1121-1142.
- [30] 阿克苏地区统计局. 2021 年阿克苏地区统计年鉴[M]. 北京:中国统计出版社,2021.
- [31] 何旭刚,买买提·沙吾提,盛艳芳,等. 基于 GEE 平台渭库绿洲棉花水分生产率遥感估算[J]. 干旱区地理,2023,46(10):1632-1642.
- [32] 祁昌贤,任燕,彭海月,等. 基于 GEE 云平台的三江源湖泊面积提取及动态变化研究[J]. 长江科学院院报,2023,40(7):1-9.
- [33] 马浩文,王永芳,郭恩亮. 基于 GEE 的翁牛特旗土地沙漠化遥感监测[J]. 干旱区研究,2023(3):1-13.
- [34] LIU J T, LI Y X, ZHANG Y, et al. Large-scale impervious surface area mapping and pattern evolution of the Yellow River delta using sentinel-1/2 on the GEE [J]. Remote Sensing, 2022, 15(1):136.
- [35] 赵安周,田新乐. 基于 GEE 平台的 1986-2021 年黄土高原植被覆盖度时空演变及影响因素[J]. 生态环境学报,2022,31(11):2124-2133.
- [36] 徐超,吕婧妤,刘昱君,等. 基于 GEE 的宝鸡峡灌区耕地灌溉面积遥感监测方法[J]. 排灌机械工程学报,2022,40(11):1167-1172.
- [37] 张祯祺,蔡惠文,张平平,等. 基于 GEE 遥感云平台的三江源植被碳源/汇时空变化研究[J]. 自然资源遥感,2023,35(1):231-242.
- [38] 宁晓刚,常文涛,王浩,等. 联合 GEE 与多源遥感数据的黑龙江流域沼泽湿地信息提取[J]. 遥感学报,2022,26(2):386-396.

(责任编辑:成纾寒)