

冯冬含,李峰,秦泉,等.基于 GEE 多特征优选的设施农业类型信息提取[J].海洋气象学报,2025,45(3):117-128.
FENG Donghan, LI Feng, QIN Quan, et al. Extraction of facility agriculture types by multi-feature selection based on GEE[J].
Journal of Marine Meteorology, 2025, 45(3): 117-128. DOI: 10.19513/j.cnki.hyqxxb.20241019001. (in Chinese)

基于 GEE 多特征优选的设施农业类型信息提取

冯冬含^{1,2}, 李峰^{1,2}, 秦泉^{1,2}, 张继波^{1,2}, 韩东枫^{1,2}, 王晗^{1,2}, 段金馈^{1,2}, 迪力努尔·热依木³

(1.山东省气象防灾减灾重点实验室, 山东 济南 250031; 2.山东省气候中心, 山东 济南 250031; 3.喀什地区气象局, 新疆 喀什 844000)

摘要 为对不同设施农业类型信息进行识别分类和精细化提取,以潍坊日光温室和塑料大棚为研究对象,基于 Google Earth Engine(GEE)云计算平台,利用 Sentinel-1 雷达卫星数据、Sentinel-2 光学卫星数据和 SRTM DEM 数据,提取光谱特征、指数特征、雷达特征、纹理特征和地形特征共 41 个特征变量并进行特征优选,其中,通过累计差法对比 14 种不同尺寸窗口,计算得到 Sentinel-2 影像最佳纹理特征。采用随机森林分类算法,根据不同特征设计 6 种组合方案开展实验,得到潍坊 2023 年 10 m 分辨率设施农业空间分布情况,探讨不同特征组合对设施农业信息提取精度的影响。结果表明:(1) Sentinel-2 提取设施农业的最佳纹理特征窗口是 27×27,最佳纹理特征组合为均值、熵、方差、相异性和对比度。(2)在光谱特征、指数特征的基础上加入纹理特征、雷达特征和地形特征对设施农业进行分类,不同类型特征对设施农业提取的有利程度由大到小依次为纹理特征、地形特征、雷达特征,与单一光谱和指数特征相比,分类精度分别提升 4.08%、1.40%、0.80%。(3)特征优选后的方案提取精度最高,设施农业的总体精度为 91.03%,Kappa 系数为 0.86。其中,塑料大棚的生产者精度和用户精度分别为 92.21%和 91.83%;日光温室的生产者精度和用户精度分别为 88.54%和 89.47%。(4)利用 Sentinel 系列影像和 SRTM DEM 地形数据,构建的潍坊多种设施农业类型遥感提取方法,可为设施农业灾害风险管理提供决策支撑。

关键词 设施农业; Sentinel 卫星; GEE; 随机森林

中图分类号: P237; S127 **文献标志码:** A **文章编号:** 2096-3599(2025)03-0117-12

DOI: 10.19513/j.cnki.hyqxxb.20241019001

Extraction of facility agriculture types by multi-feature selection based on GEE

FENG Donghan^{1,2}, LI Feng^{1,2}, QIN Quan^{1,2}, ZHANG Jibo^{1,2}, HAN Dongfeng^{1,2},
WANG Han^{1,2}, DUAN Jinkui^{1,2}, Dilinuer Reyimu³

(1. Key Laboratory for Meteorological Disaster Prevention and Mitigation of Shandong, Jinan 250031, China; 2. Shandong Climate Center, Jinan 250031, China; 3. Kashgar Meteorological Service, Kashgar 844000, China)

Abstract To distinguish and precisely extract the information of different types of facility agriculture,

收稿日期: 2024-10-19; 修回日期: 2025-02-22

基金项目: 山东省自然科学基金项目(ZR2020MF130, ZR202211260263); 山东省气象局科研项目(2023sdqx12, 2021sdqx03); 山东省气候中心创新项目(2023QHXC04)

第一作者: 冯冬含, dorisfdh@163.com。

通信作者: 秦泉, qqquan0303@163.com。

this study takes the solar greenhouses and polyhouses in Weifang as the research objects using the Google Earth Engine (GEE) platform. By integrating Sentinel-1 radar data, Sentinel-2 optical data and SRTM DEM data, 41 feature variables are extracted, including the features of spectrum, index, radar, texture and topography, and the optimal selection of features is conducted. By comparing 14 window scales with the cumulative difference method, the optimal texture features for Sentinel-2 imagery are identified. The experiments with 6 combinations of features are designed using the random forest classification algorithm to generate a distribution map of facility agriculture in Weifang with the resolution of 10 m in 2023, and the impacts of different combinations on the accuracy of extraction are explored. The results are outlined below. (1) The optimal window scale of the texture feature for extracting facility agriculture using Sentinel-2 is 27×27 , and the best combination of texture features includes Mean, Entropy, Variance, Dissimilarity and Contrast. (2) Besides the features of spectrum and index, the features of texture, radar and topography are used to classify facility agriculture. The order of different features beneficial to extract facility agriculture from high to low is texture, topography and radar, and compared to the classification using only spectrum and index, the accuracy increases by 4.08%, 1.40% and 0.80%, respectively. (3) The optimal selection of features yields the highest classification accuracy, with an overall accuracy of 91.03% and a Kappa coefficient of 0.86. The producer's accuracy and user's accuracy for polyhouses are 92.21% and 91.83%, respectively, while those for solar greenhouses are 88.54% and 89.47%, respectively. (4) The remote sensing extraction method for various facility agriculture types in Weifang using Sentinel imagery and SRTM DEM topography data offers decision-making support for disaster risk management of facility agriculture.

Keywords facility agriculture; Sentinel satellite; GEE (Google Earth Engine); random forest

0 引言

设施农业是现代农业的重要组成部分,作为一种显著延长作物生长季节和提高作物产量的新型生产方式,能够有效提高资源利用率并增强农业系统的防灾减灾能力,是保障“菜篮子工程”顺利开展的重要技术措施^[1]。近年来气候变化导致极端天气事件频发^[2-3],涝灾、风灾等气象灾害风险的加剧直接威胁设施农业生产^[4-5]。由于建筑结构、材料的差异,不同类型的设施农业对灾害的防御能力完全不同。因此,及时准确地获取设施农业分布情况,可为防灾减灾提供决策支持,并且在促进农业产业结构调整和提高城乡居民的生活水平等方面具有十分重要的现实意义。

随着农业经济、信息技术的迅速发展,设施农业的占地面积逐年攀升。依靠人工实地调查设施农业不仅费时费力,而且很难及时更新大范围设施农业的空间分布数据。与传统地面调查方法相比,遥感技术提供了更高效的手段,并具有成本低、效率高、大范围同步观测的优势^[6-7],在设施农业的信息提取方面提供了新的思路。目前已有研究尝试使用遥感数据获取设施农业信息,并取得良好进展。提取设施农业通

常有2类方法:一是基于像元的支持向量机(support vector machines, SVM)^[8]、随机森林(random forest, RF)和神经网络(neural network, NN)^[9]等机器学习分类方法,其优点是计算效率高且适用性广泛,冯权泷等^[10]利用随机森林模型获取全国塑料大棚分布数据,平均精度为87.45%;二是基于影像分割的面向对象分类方法,例如多尺度分割、灰度分割、知识分割等,可以保证地物在空间上的连续性,但参数依赖性强,仅适合小区域提取,例如吴锦玉等^[11]、Hasituya等^[12]均选择县级区域作为研究区。

现有的研究主要集中在设施农业的整体提取上,但在大范围内针对不同设施农业类型信息提取的方法研究较少。随着遥感技术不断发展,中高分辨率卫星逐步发射,可为地物分类提供更多细节信息。欧洲航天局(European Space Agency, ESA)发布的Sentinel-2光学影像数据空间分辨率达10 m,满足设施农业类型提取需求^[13]。然而影像的空间分辨率越高,同一地物内的异质性越强,从而影响分类精度^[14]。因此除Sentinel-2的光谱特征外,还需考虑其他特征参与精细化分类研究。Sentinel-1搭载的合成孔径雷达(synthetic aperture radar, SAR)可以全天候获取地表信息,提供的后向散射系数能够帮助识别设

施农业的纹理特征。融合光学影像和 SAR 影像,能够实现多源遥感数据互补^[15-16],进一步提升识别和分类精度。但长时间序列、大空间尺度的遥感数据存储和计算等处理对软硬件环境提出了更高的要求。谷歌于 2015 年发布了专门用于处理遥感影像的 Google Earth Engine (GEE) 云计算平台,能够高效管理和处理海量遥感影像,为提取设施农业信息提供了强有力的技术支撑。相比于 ERDAS、ENVI 等传统遥感图像处理软件,云端处理影像能减少过程数据的生成,用户可以快速实现遥感影像批量处理和结果可视化等功能^[17],极大地提高了数据处理效率。基于此,国内外学者开展的一系列研究^[18-20]都表明 GEE 云计算平台在遥感数据分析方面有着巨大的应用潜力。

以山东潍坊为研究区,利用 GEE 云计算平台,将 Sentinel-1 雷达影像和 Sentinel-2 光学影像作为数据源,构建光谱、指数、雷达、纹理和地形特征集合,设计 6 种设施农业(日光温室、塑料大棚)提取方案,采用

随机森林算法分类并进行精度评价,筛选出最佳特征组合,实现大范围提取设施农业类型信息,为设施农业精细化分类提供新思路。此研究可对潍坊设施农业的产业发展和防灾减灾提供重要的决策支持。

1 数据与方法

1.1 研究区概况

潍坊位于山东半岛的西部(35.70°~37.29°N, 118.17°~120.01°E;图 1),属暖温带半湿润大陆性季风气候,降水、光照资源丰富^[21],适合多种农作物生长。地形以平原为主,兼有丘陵和山地,地势自南向北逐渐降低。南依泰沂山脉,中部为潍北冲积平原,北临渤海莱州湾,地势平坦、土地肥沃。潍坊是中国北方地区最大的蔬菜生产和集散地,设施农业面积位居全省第一。在平原地区聚集不同类型设施农业,分布最为广泛的是日光温室和塑料大棚,设施主要用于种植高附加值的经济作物,如番茄、黄瓜、辣椒、西瓜、草莓等。

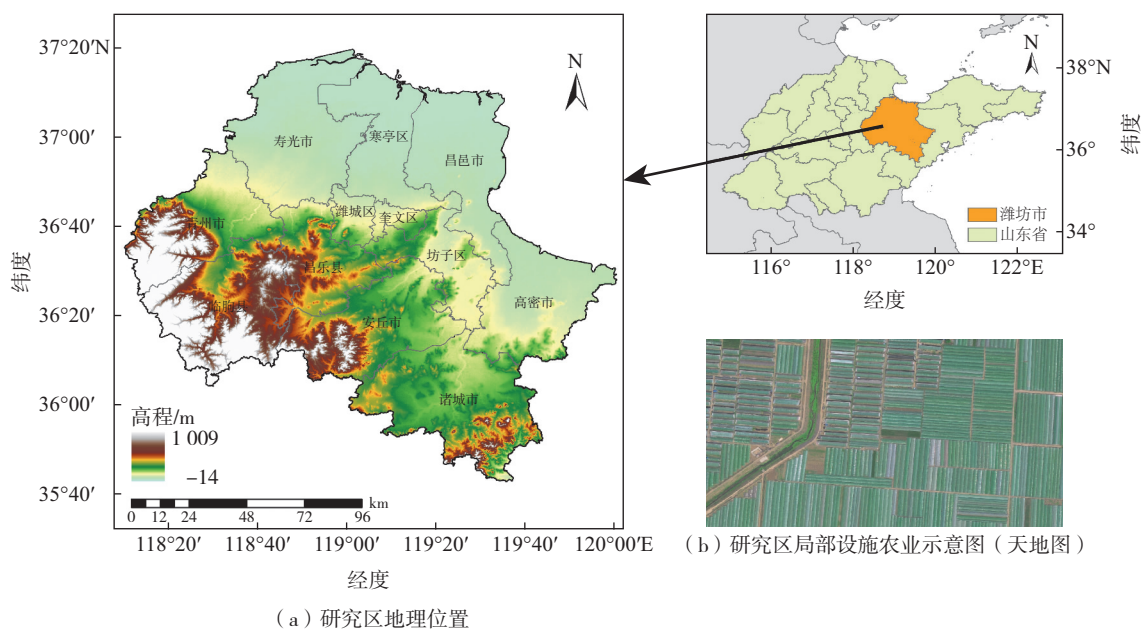


图 1 研究区地理位置及局部设施农业示意图

Fig.1 Geographical location of the study area and sketch map of facility agriculture in local areas

1.2 数据

1.2.1 遥感数据获取与预处理

研究使用高空间分辨率的 Sentinel-1 和 Sentinel-2 卫星影像,数据源自哥白尼数据空间生态系统(Copernicus Data Space Ecosystem, <https://dataspace.copernicus.eu>)。Sentinel-1 选用干涉宽幅(interferometric wide-swath, IW)模式的地面分辨距

离(ground resolving distance, GRD)数据,极化方式选择适用于地表植被特征监测的 VV + VH^[22]。Sentinel-2 影像数据涵盖从可见光、近红外到短波红外共 13 个光谱波段,各波段具有不同的空间分辨率(10 m、20 m 和 60 m),为提高数据一致性,对所有波段重采样至 10 m 分辨率。实验收集成像时间范围为 2023 年 1—12 月的卫星影像(表 1)。为减少

云层遮挡对提取结果的影响, 优先筛选云量小于 10% 的 Sentinel-2 影像数据, 并利用 QA60 波段进行云掩膜。其中, 8 月和 10 月云量过多, 无法直接使用单景或多景拼接影像, 因此采用中值合成法得到月尺度无云合成影像。

表 1 影像数据采集日期		
Table 1 Acquisition date of imagery data		
月份	数据采集日期	
	Sentinel-1	Sentinel-2
1	2023 年 1 月 20 日	2023 年 1 月 21 日
2	2023 年 2 月 13 日	2023 年 2 月 20 日
3	2023 年 3 月 9 日	2023 年 3 月 7 日
4	2023 年 4 月 14 日	2023 年 4 月 16 日
5	2023 年 5 月 8 日	2023 年 5 月 1 日
6	2023 年 6 月 1 日	2023 年 6 月 15 日
7	2023 年 7 月 31 日	2023 年 7 月 5 日
8	2023 年 8 月 24 日	2023 年 8 月
9	—	2023 年 9 月 28 日
10	2023 年 10 月 11 日	2023 年 10 月
11	2023 年 11 月 16 日	2023 年 11 月 12 日
12	—	2023 年 12 月 7 日

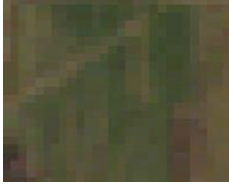
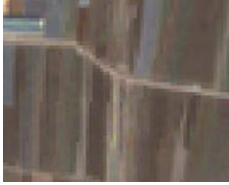
1.2.2 地形数据

研究区地形包括平原、丘陵和山地, 加入地形数

据的目的是拟提高分类精度。实验基于 GEE 平台调用美国地质调查局 (United States Geological Survey, USGS) 提供的 30 m 空间分辨率的数字高程模型 (digital elevation model, DEM) 数据产品 (USGS/SRTMGL1_003), 重采样至 10 m 后按研究区边界裁剪, 使其与 Sentinel 数据分辨率匹配, 便于多源数据集成与对比分析。

1.2.3 样本采集

在设施农业信息提取过程中, 样本类别的划分和选取至关重要。结合实地调查和山东省地理信息公共服务平台天地图 2 m 历史影像 (下载地址为 <https://shandong.tianditu.gov.cn>) 进行目视解译, 建立研究区内日光温室、塑料大棚、耕地、道路、水体、森林、建设用地、裸地等主要地物类型的遥感解译标志 (表 2)。基于潍坊矢量边界, 根据解译标志分散随机绘制 2 800 个样本, 包括日光温室和塑料大棚各 580 个, 其他地物 1 640 个。按 7:3 比例划分为训练样本和验证样本, 即 1 960 个训练样本和 840 个验证样本 (图 2)。

表 2 地物类型解译标志					
Table 2 Explanation of symbols of land cover types					
地物类型	Sentinel-2 真彩色影像	解译标志说明	地物类型	Sentinel-2 真彩色影像	解译标志说明
温室		多为横向排列, 浅蓝色、部分白色	水体		大面积分布, 蓝黑色
塑料大棚		多为纵向排列, 浅蓝色	森林		多分布于山区
耕地		规则块状、成片分布, 内部多为矩形, 绿色或浅绿色	建设用地		块状几何特征明显、边界清晰
道路		条带状, 浅黄色	裸地		规则块状、成片分布, 内部多为矩形, 黄色为主

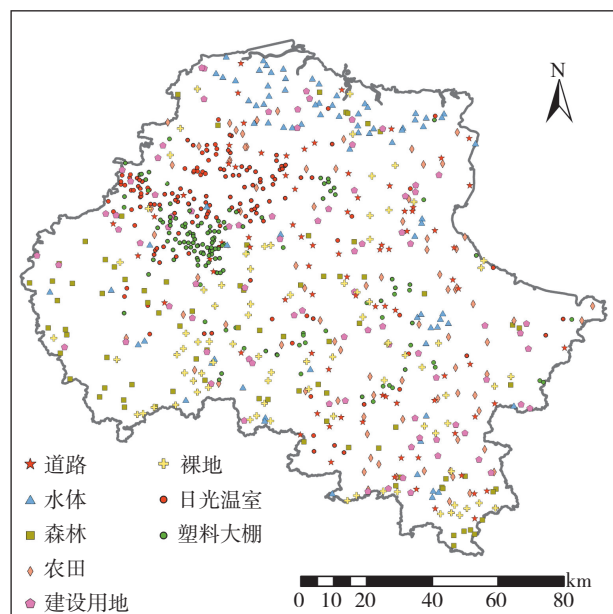


图2 潍坊不同地物类型验证样本的分布
Fig.2 Distribution of test samples for different land cover types in Weifang

1.3 研究方法

1.3.1 技术路线

研究技术路线如图3所示。使用 GEE 云计算平台,快速获取覆盖潍坊的 Sentinel 遥感影像和 DEM 数据,并进行重采样、去云、裁剪等处理。在此基础上,提取光谱特征、指数特征、地形特征、雷达特

征和纹理特征构建特征集。基于不同特征组合,设计6种设施农业提取方案:1.使用光谱特征和指数特征作为基础方案,用以评估其独立的分类性能;2.加入地形特征,提供额外的空间位置信息,有助于提高分类模型的判别能力;3.在基础方案中加入雷达特征用于反映地物表面粗糙度、电磁波散射特性等属性,是对光学影像的重要补充;4.引入基于 Sentinel-2 影像计算的纹理特征,捕捉地物结构的细微差异,从而提升设施农业精细化分类能力;5.综合全部特征,全面利用多源数据的优势最大化提升分类精度;6.在全部特征组合的基础上,通过特征优选方法筛选贡献度高的变量,剔除冗余信息,进一步提升分类效率和精度。所有方案均采用随机森林监督分类方法进行提取,并对分类结果进行精度评价,选择精度最高的特征组合方案,获取潍坊设施农业空间分布信息。

1.3.2 特征选取

(1)光谱特征。在 Sentinel-2 光学影像的 13 个原始波段中,剔除 B1、B10 这 2 个无关波段,选取其他 11 个光谱波段作为原始波段特征,并计算 19 个常用遥感指数(表3),如归一化植被指数(NDVI)、归一化水体指数(NDWI)、红边位置指数(REPI)、土壤调节植被指数(SAVI)等。计算结果在实验中作为独立的指数特征波段,与原始波段并列参与分类任务。

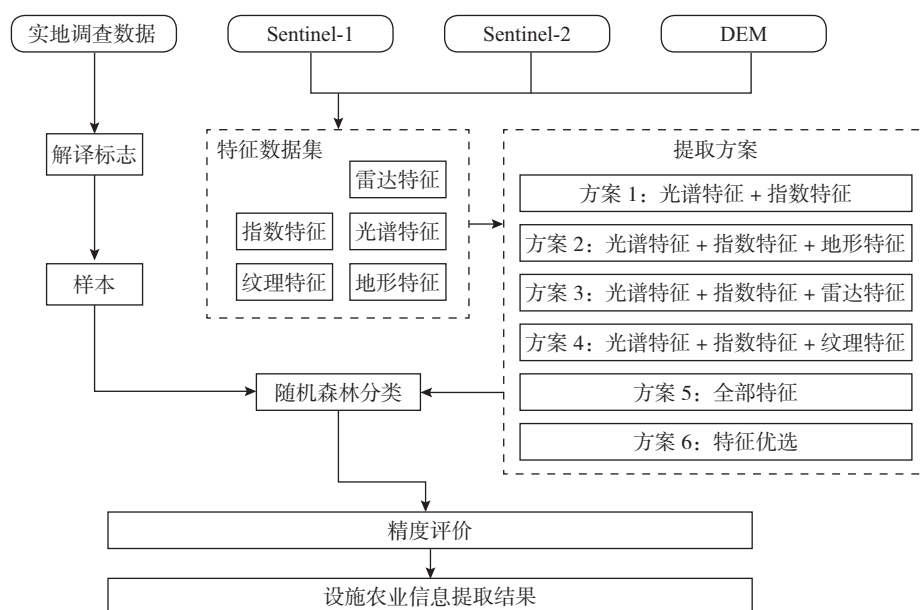


图3 技术路线
Fig.3 Technical route diagram

(2) 雷达特征。Sentinel-1 有 4 种成像模式(干涉宽幅模式 IW、扩展宽幅模式 EW、波束模式 WV、条带模式 SM)和 2 类极化方式,分为单极化方式(HH 或 VV)和双极化方式(HH+HV 或 VV+VH)。其中,IW 模式广泛应用于陆地表面的遥感研究^[23],特别是 VV 和 VH 偏振组合,在地表特征提取和分析中表现出优异的效果。本研究选用 VV+VH 作为独立的雷达特征波段参与特征构建。

(3) 纹理特征。纹理特征是地物提取和识别中的重要参数,用来描述图像中像素灰度值的空间分布,这种局部的统计信息能够帮助提取复杂地物的特征。常用的纹理分析方法是灰度共生矩阵(GLCM)^[24],因其计算过程简单、结构信息丰富,故在遥感提取研究中得到广泛应用。GLCM 算法需要 8 位灰度图像作为输入,因此此实验通过原始影像的红(RED)、绿(GREEN)、蓝(BLUE)波段的线性组合生成所需灰度图像(gray),计算公式如下。

$gray = (0.3 \times RED) + (0.59 \times GREEN) + (0.11 \times BLUE)$ (1)

实验选取 8 种主要的纹理特征,包括角二阶矩(ASM)、相关性(Correlation)、对比度(Contrast)、熵(Entropy)、均值(Mean)、方差(Variance)、相异性(Dissimilarity)、同质性(Homogeneity)。

(4) 地形特征。日光温室和塑料大棚的建造受地形限制,加入地形特征尝试提高分类精度。除海拔高度(Elevation)外,增加 3 种常用的地形特征,即坡度(Slope)、坡向(Aspect)、山体阴影(Hillshade),这 4 种特征将作为独立波段参与特征构建。

1.3.3 Jeffries-Matusita 距离

Jeffries-Matusita(JM)距离^[25]常用于遥感影像分类中评估不同类别之间的可分性。将全年不同月份的 JM 距离进行比较,JM 距离越大,表明设施农业与其他地物类别的可分性越好,在此研究中 JM 距离最大的月份被认为是提取的最佳时相。

表 3 参与分类的特征
Table 3 Features involved in classification

类别	分类特征	简称	类别	分类特征	简称
光谱特征	蓝光波段	B2	雷达特征	垂直水平极化	VH
	绿光波段	B3		垂直垂直极化	VV
	红光波段	B4	指数特征	叶绿素绿光植被指数	CIg
	红边 1 波段	B5		红边叶绿素指数	CIre
	红边 2 波段	B6		增强植被指数	EVI
	红边 3 波段	B7		绿度指数	GI
	近红外波段	B8		地表水分指数(B8、B11)	LSWI1
	红边 4 波段	B8A		地表水分指数(B8、B12)	LSWI2
	水汽波段	B9		改进的归一化水体指数	MNDWI
	短波红外 1 波段	B11		改进的土壤调节植被指数	MSAVI
	短波红外 2 波段	B12		地面叶绿素指数	MTCI
地形特征	坡向	Aspect		归一化建筑指数	NDBI
	海拔高度	Elevation		归一化差异红边植被指数	NDRE
	山体阴影	Hillshade		归一化植被指数	NDVI
	坡度	Slope		归一化水体指数	NDWI
纹理特征	角二阶矩	ASM		新型植被指数	NIRv
	相关性	Correlation		红边位置指数	REPI
	对比度	Contrast		比值植被指数	RVI
	相异性	Dissimilarity		土壤调节植被指数	SAVI
	熵	Entropy		转换型土壤调节植被指数	TSAVI
	均值	Mean		水体增强指数	WRI
	方差	Variance			
	同质性	Homogeneity			

设 μ_i 和 μ_j 分别为类别*i*和*j*的均值向量, Σ_i 和 Σ_j 分别为对应的协方差矩阵。JM 距离的定义如下。

$$JM_{ij} = \sqrt{2 \left\{ 1 - \exp \left[-\frac{1}{8} (\mu_i - \mu_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j) - \frac{1}{2} \ln \left(\frac{\frac{\Sigma_i + \Sigma_j}{2}}{\sqrt{|\Sigma_i| |\Sigma_j|}} \right) \right] \right\}}$$

(2)

其中: $\mu_i - \mu_j$ 表示类别均值向量的差异, $\frac{\sum i + \sum j}{2}$ 表示类别协方差矩阵的均值。

1.3.4 累计差

在计算纹理特征时,滑动窗口大小(邻域的尺寸)对分类结果有显著影响。较小窗口容易引入过多噪声,导致分类结果不稳定。而较大窗口会模糊图像中的细节特征,降低分类精度。因此,选择合适的窗口大小是获取纹理特征的第一步。此研究采用累计差(Δf)^[26]衡量地物区分度以确定窗口大小,累计差越大,表明地物在当前窗口下的区分度越高,其计算公式如下。

$$\Delta f = \sum_{i=1}^k (|u_i - u_{i+1}| + |u_i - u_{i+2}| + \dots + |u_i - u_{i+k-1}|) \quad (3)$$

其中: Δf 代表某一纹理特征变量的累计差,此研究共有8个纹理特征,因此 $f=1, 2, \dots, 8$; k 为地物类别总数; u_i 表示第 i 类地物在某特征变量上的均值。

1.3.5 分类方法

RF 算法^[27]是一种基于集成学习的监督分类算法,其核心思想是通过构建大量相互独立的决策树来提高分类的准确性和稳定性,具有良好防止过拟合能力。相比其他分类器,RF 在处理高维遥感影像数据时表现出更好的性能,对噪声和缺失数据具有较强的鲁棒性,适用于多类别和多特征的遥感分类^[28]。

在 GEE 云计算平台中,使用 RF 分类算法时需要设置关键参数,此研究设置决策树的数量为100个,每个决策节点使用的变量数为特征总数的平方根。

1.3.6 精度评价

在遥感图像分类中,混淆矩阵是评估分类模型性能的常用工具,用于表示实际类别与模型分类结果的对比情况。此研究通过计算分类结果的混淆矩阵对6种特征组合方案进行精度评价,评价指标包括总体精度(overall accuracy, OA)、Kappa系数、用户精度(user's accuracy, UA)和生产者精度(producer's accuracy, PA)。

2 结果与分析

2.1 最佳时像选取

根据目视解译的样本,使用 GEE 云计算平台,对 Sentinel-2 影像不同月份下设施农业与其他地物的 JM 距离进行计算,确定提取的最佳时相。如图 4

所示,2023 年 Sentinel-2 影像塑料大棚与其他地物的 JM 距离均大于日光温室与其他地物的 JM 距离,且二者在全年变化趋势相似,夏季 JM 距离小,冬季 JM 距离大,2 月 20 日均达到最大。这表明设施农业在 2 月的光谱差异性最大,更容易与其他地物区分。因此,选取 2 月 20 日的 Sentinel 影像进行设施农业提取。

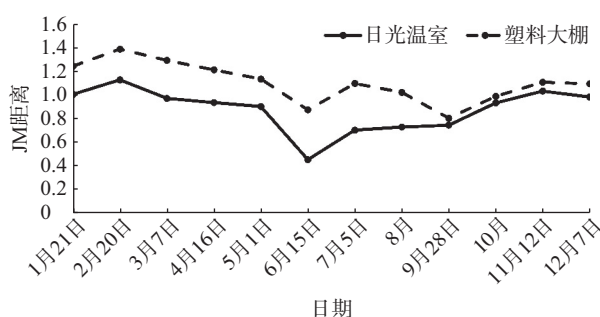


图4 不同日期下设施农业与其他地物 JM 距离
Fig.4 JM distance between facility agriculture and other land cover types on different dates

2.2 纹理特征提取

GEE 云计算平台提供的 GLCM 算法中可定义计算窗口大小。研究选取 3×3 、 7×7 、 11×11 、 15×15 、 19×19 、 23×23 、 27×27 、 31×31 、 35×35 、 39×39 、 43×43 、 47×47 、 51×51 、 55×55 共 14 种尺寸的滑动窗口,在最佳提取时相的 Sentinel-2 灰度图像上计算不同窗口大小下各地物(日光温室、塑料大棚及其他地物)的纹理特征累计差。如图 5a 所示,累计差随着窗口的增大,先增加后减小,在 27×27 窗口下累计差达最大,此时设施农业与其他地物的纹理特征区分度最高,选择 27×27 窗口为最佳提取窗口。图 5b 展示了最佳提取窗口下各纹理特征的累计差,其中方差特征的累计差最高,熵、相异性、对比度、均值依次降低,同质性、角二阶矩、相关性的累计差最低且与前 5 个纹理变量相差过大,累计差均小于 0.5。因此选择方差、熵、相异性、对比度、均值作为纹理特征参与模型训练。

2.3 特征对分类的贡献分析

为了筛选出对分类精度贡献较大的特征,避免不重要的噪声特征干扰,使用基于基尼不纯度的特征重要性评估,实验共获取特征变量 41 个。由图 6 可以直观看出,地形特征中的 Elevation 重要性最高,这反映出设施农业的分布可能与地形起伏密切

相关,日光温室和塑料大棚多修建于平坦的地理条件下,与潍坊的实际情况相符合。后面依次是均值、方差、相异性、熵,表明纹理特征在此次分类重要性很高,这对于捕捉设施农业的空间结构具有重要作用。此外,雷达波段 VV、VH 和光谱波段 B9、B12 也

具有较高的重要性。光谱特征尤其是与近红外和可见光相关的波段常用于植被监测,在提取与设施农业相关的植被信息时起到了关键作用。特征重要性排序依次为:纹理特征、光谱特征、雷达特征、指数特征、地形特征。

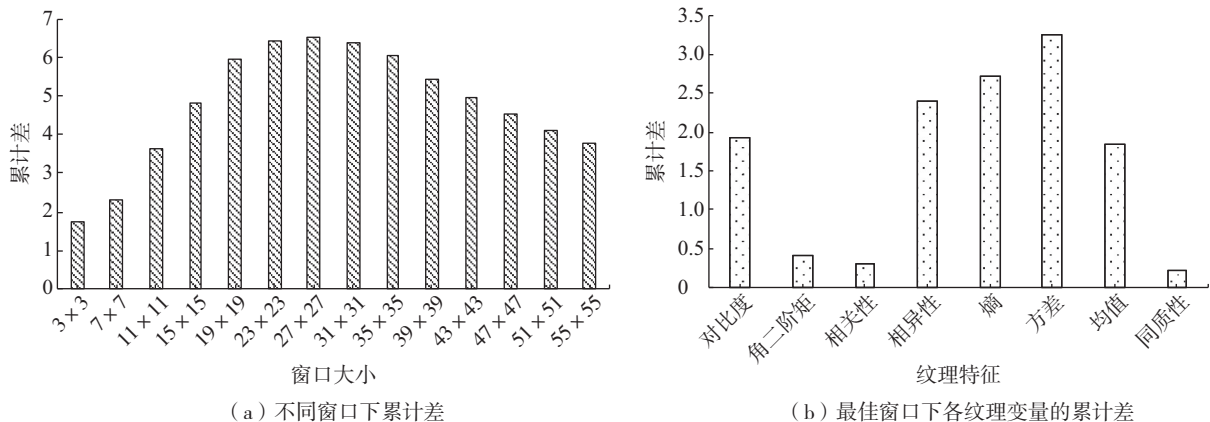


图 5 不同窗口下累计差与最佳窗口下各纹理变量的累计差

Fig.5 Cumulative difference under different windows and cumulative difference of texture variables under the optimal window

特征冗余不仅会增加计算复杂度,还会出现过拟合的风险,导致分类精度降低,因此需要对特征进行筛选。根据重要性顺序依次增加特征训练 RF 模型,计算

每个模型分类结果的 OA。由图 6 可以看到,添加 NDVI 后,整体分类精度趋于平稳且无显著提升,因此保留包括 NDVI 在内的前 26 个特征构成最优特征集。

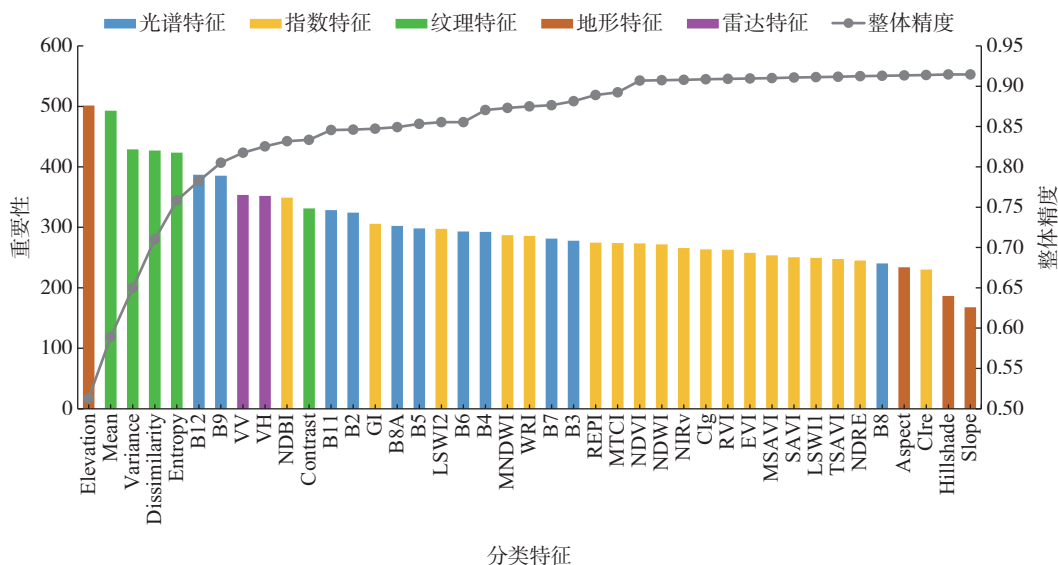


图 6 特征重要性排序

Fig.6 Ranking of feature importance

2.4 精度评价

使用固定种子的 RF 算法对 6 种特征组合方案进行分类,结果如表 4 所示。方案 1 仅包括光谱和指数特征,分类精度在各方案中最低。尽管光谱信息和植被指数是常用的遥感特征,但对于复杂的设

施农业,并不足以完全反映其特征差异。在方案 1 基础上分别加入地形特征、雷达特征、纹理特征后,分类精度均有提高。其中,雷达特征对分类精度的提升最少,塑料大棚的 UA 提高 2.28%,而日光温室仅提高 0.54%,说明雷达特征对日光温室提取精度

的贡献有限。在方案 1 中增加纹理特征后,OA 提升至 88.78%,且日光温室和塑料大棚的分类精度均有大幅提升。对比方案 2 和方案 3,纹理特征对分类精度的提高最显著,表明纹理特征能够捕捉影像中局部空间关系和细节,对于区分形态复杂的设施农业尤其有效。方案 5 融合所有特征,OA 达 89.38%,经过优选特征后,方案 6 的 OA 相较方案 1 提高了 6.33%,Kappa 系数达 0.86,有效减少特征冗余,进一步提高了模型的鲁棒性和分类精度,在所有方案中取得最好的分类效果。根据 OA 和 Kappa 系数,不同特征对设施农业提取的有利程度排序为:纹理特征、地形特征、雷达特征、光谱特征。

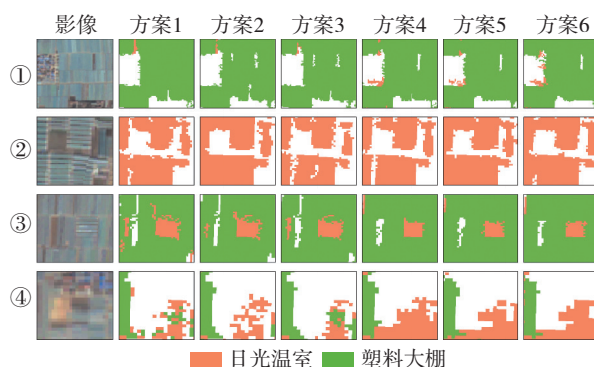
表 4 不同方案下设施农业分类精度

Table 4 Classification accuracy of facility agriculture under different schemes

方案	地物信息	UA/%	PA/%	Kappa 系数	OA/%
1	日光温室	80.00	78.31	0.76	84.70
	塑料大棚	84.87	86.13		
	其他地物	86.88	89.09		
2	日光温室	82.06	80.32	0.79	86.10
	塑料大棚	87.16	89.14		
	其他地物	88.03	87.69		
3	日光温室	80.54	79.68	0.78	85.50
	塑料大棚	87.15	87.96		
	其他地物	87.59	87.59		
4	日光温室	87.02	85.63	0.83	88.78
	塑料大棚	90.68	91.42		
	其他地物	88.18	88.56		
5	日光温室	87.95	86.59	0.84	89.38
	塑料大棚	90.63	91.81		
	其他地物	89.30	89.30		
6	日光温室	89.47	88.54	0.86	91.03
	塑料大棚	91.84	92.21		
	其他地物	91.45	91.84		

基于 GEE 云计算平台和随机森林算法,提取出潍坊 2023 年设施农业的分布区域。图 7 展示了日光温室和塑料大棚集中分布区域、交错分布区域、复杂背景区域等 4 个典型区域的提取情况。对于单一类型且大面积分布的设施农业,所有方案的分类结果差异较小。然而,在交错分布和复杂背景区域,6 种方案表现出较大差异。方案 1—3 均可识别设施农业的基本轮廓,但在边缘处出现明显的分类错误和噪点,导致边界不清晰、误分类现象较多。方案 4 纹理特征的加入显著提高了分类精度,特别是在细节处,混合像元更多的被分为其他地物,使道路、农田的边界更

加清晰,但会使设施农业统计面积减少。方案 5 综合了所有特征,包括光谱特征、指数特征、地形特征、雷达特征和纹理特征。虽然该方案理论上应达到最优效果,但实际分类结果并未明显优于方案 4。这可能是特征过多导致信息冗余,引入了额外的噪声。而特征优选后的方案 6 在不同区域下均能够较好地捕捉设施农业的完整性和边界特征。



①—塑料大棚集中分布区域; ②—日光温室集中分布区域;
③—日光温室和塑料大棚交错分布区域; ④—复杂背景区域。

图 7 不同方案设施农业提取细节
Fig.7 Details of facility agriculture extraction under different schemes

图 8 展示了特征优选方案的设施农业提取结果。可以看出,潍坊的设施农业主要集中在寿光和青州,总面积占全市设施农业面积的 66%(表 5),但日光温室和塑料大棚的空间分布存在明显差异。日光温室主要集中分布在寿光、青州、临朐等区域,其中寿光面积达 40.36 万亩^①,占全市日光温室总面积的 51.45%,是日光温室设施农业发展的核心区域。其次为青州,面积为 17.09 万亩,占 21.79%。与日光温室相比,塑料大棚的分布更分散。青州和昌乐的塑料大棚面积较大,分别为 16.89 万亩和 8.79 万亩,合计占全市塑料大棚总面积的 57.52%。

与地形图叠加发现,设施农业主要建在地势平坦的平原地区,极少建立在山地和丘陵地区。从地貌的角度来看,寿光南部和青州东北部地势相对平坦,主要为冲积平原,土地开阔平整,减少了建设温室大棚基础设施的成本,有利于大规模设施农业的布局和管理^[29]。在人文因素方面,寿光作为“中国蔬菜之乡”,拥有丰富的设施农业经验和完善的产业链条。青州北部和昌乐北部由于与寿光接壤,很大程度上受其技术和市场的辐射和影响,形成了设施农业集聚区。

①1 亩 $\approx 0.0667 \text{ hm}^2$ 。

得了较高的精度,但研究中仍有不足的地方。首先,此研究主要集中在单一时相影像的提取,未充分考虑多时相数据对提取精度的影响,进一步的研究可以结合设施农业内部种植作物的种类,加入多时相数据,更好地捕捉季节性变化和设施农业使用状态。其次,随着人工智能的不断发展,深度学习技术在许多领域表现出优异的性能,未来研究可以探索利用深度学习算法,实现大面积、更加精细的设施农业信息提取。

4 小结

(1)根据样本计算不同设施农业类型与其他地物的 JM 距离,得出日光温室和塑料大棚的最佳提取时相均为 2 月。

(2)使用累计差法从 14 种窗口大小中确定提取设施农业的最佳纹理特征窗口是 27×27 ,最佳纹理特征组合为均值、熵、方差、相异性和对比度。

(3)不同的特征组合方案对提取结果影响不同。分类特征对设施农业的重要性排序依次为:纹理特征、地形特征、雷达特征。3 个特征的加入使得分类精度分别提升 4.81%、1.66%、0.94%。

(4)6 个方案中特征优选方案提取不同设施农业类型信息效果最佳,OA 为 91.03%,Kappa 系数为 0.86。其中,塑料大棚的 PA 和 UA 分别为 92.21% 和 91.84%;日光温室的 PA 和 UA 分别为 88.54% 和 89.47%。这表明该方法能够有效获取不同设施农业类型的空间分布和面积信息。

参考文献:

- [1] 杨婧暄.我国农业大棚现状及未来发展趋势[J].种子科技,2023,41(14):142-144.
- [2] 刁秀广.强冰雹和强降水超级单体风暴双偏振特征差异性[J].海洋气象学报,2024,44(3):23-36.
- [3] 侯淑梅,唐巧玲,史茜,等.2020 年 5 月山东强对流天气特点及成因[J].海洋气象学报,2023,43(4):47-61.
- [4] 谭方颖,王建林,张蕾,等.山东省设施农业火灾风险评估研究[J].中国生态农业学报(中英文),2024,32(1):174-182.
- [5] 魏瑞江,王鑫.国内设施种植气象灾害研究进展[J].海洋气象学报,2021,41(4):73-81.
- [6] 李莉,刘志红,韩晨琛,等.基于卫星遥感的临沂市冬小麦种植时空变化特征分析[J].海洋气象学报,2020,40(3):126-135.
- [7] 段金馈,李峰,秦泉,等.高分六号红边波段对夏玉米种植区识别的影响[J].海洋气象学报,2023,43(3):104-112.
- [8] KOC-SAN D. Evaluation of different classification techniques for the detection of glass and plastic greenhouses from WorldView-2 satellite imagery[J]. J Appl Remote Sens,2013,7(1):073553.
- [9] 付利钊,杨青岗,陈永立,等.基于 SFNet-F 地物识别技术的农业大棚信息提取[J].测绘通报,2024(7):65-70.
- [10] 冯权泷,牛博文,朱德海,等.2019 年全国农业塑料大棚遥感分类数据集[J].中国科学数据(中英文网络版),2021,6(4):153-170.
- [11] 吴锦玉,刘晓龙,柏延臣,等.基于 GF-2 数据结合多纹理特征的塑料大棚识别[J].农业工程学报,2019,35(12):173-183.
- [12] HASITUYA, CHEN Z X, WANG L M, et al. Monitoring plastic-mulched farmland by Landsat-8 OLI imagery using spectral and textural features[J]. Remote Sens,2016,8(4):353.
- [13] 刘天宇,赵展,史同广.一种基于 Sentinel-2 的塑料大棚提取方法[J].农业工程,2021,11(10):91-98.
- [14] TANG G, WANG X Y, ZHAO H W, et al. Attention in attention for hyperspectral with high spatial resolution (H) image classification[J]. IEEE Trans Geosci Remote Sens,2023,62:5501113.
- [15] 乔树亭,叶回春,黄文江,等.基于 Sentinel-1/2 影像的水稻种植面积提取方法研究:以三江平原为例[J].遥感技术与应用,2023,38(1):78-89.
- [16] HAFNER S, NASCETTI A, AZIZPOUR H, et al. Sentinel-1 and Sentinel-2 data fusion for urban change detection using a dual stream U-Net[J]. IEEE Geosci Remote Sens Lett,2022,19:4019805.
- [17] 王小娜,田金炎,李小娟,等.Google Earth Engine 云平台对遥感发展的改变[J].遥感学报,2022,26(2):299-309.
- [18] GANJIRAD M, BAGHERI H. Google Earth Engine-based mapping of land use and land cover for weather forecast models using Landsat 8 imagery[J]. Ecol Inform,2024,80:102498.
- [19] KAZEMI GARAJEH M, HAJI F, TOHIDFAR M, et al.

- Spatiotemporal monitoring of climate change impacts on water resources using an integrated approach of remote sensing and Google Earth Engine [J]. *Sci Rep*, 2024, 14(1):5469.
- [20] 韩东枫,李峰,秦泉,等.基于 GEE 和 Sentinel-1/2 数据的夏玉米种植面积精细化识别方法[J].*海洋气象学报*,2024,44(3):122-132.
- [21] LI C M, DAI Z X, YANG L N, et al. Spatiotemporal characteristics of air quality across Weifang from 2014–2018[J]. *Int J Environ Res Public Health*, 2019, 16(17):3122.
- [22] 孙政,周清波,杨鹏,等.基于星载极化 SAR 数据的农作物分类识别进展评述[J].*中国农业资源与区划*,2019, 40(11):63-71.
- [23] MA Y Y, WANG Z M, ZHANG B J, et al. The spatiotemporal surface velocity variations and analysis of the Amery Ice Shelf from 2000 to 2022, East Antarctica [J]. *Remote Sens*, 2024, 16(17):3255.
- [24] HARALICK R M, SHANMUGAM K, DINSTEN I. Textural features for image classification[J]. *IEEE Trans Syst Man Cybern*, 1973, SMC-3(6):610-621.
- [25] BRUZZONE L, ROLI F, SERPICO S B. An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection[J]. *IEEE Trans Geosci Remote Sens*, 1995, 33(6):1318-1321.
- [26] 姚新华,金佳,徐飞飞,等.太湖流域果树提取的光谱和纹理特征选择研究[J].*中国生态农业学报(中英文)*, 2019, 27(10):1596-1606.
- [27] ZHAO Y F, ZHU W W, WEI P P, et al. Classification of Zambian grasslands using random forest feature importance selection during the optimal phenological period[J]. *Ecol Indic*, 2022, 135:108529.
- [28] 侯蒙京,殷建鹏,葛静,等.基于随机森林的高寒湿地地区土地覆盖遥感分类方法[J].*农业机械学报*,2020, 51(7):220-227.
- [29] GUO B, ZHOU B, ZHANG Z M, et al. A critical review of the status of current greenhouse technology in China and development prospects[J]. *Appl Sci*, 2024, 14(13):5952.
- [30] 朱德海,刘逸铭,冯权洸,等.基于 GEE 的山东省近 30 年农业大棚时空动态变化研究[J].*农业机械学报*, 2020, 51(1):168-175.
- [31] 汤紫霞,李蒙蒙,汪小钦,等.基于 GF-2 遥感影像的葡萄大棚信息提取[J].*中国农业科技导报*, 2020, 22(11):95-105.
- [32] 王乌云,李斐,哈斯图亚,等.面向对象多层次多尺度分割的农用大棚类型信息提取[J].*中国农业大学学报*, 2024, 29(8):223-236.
- [33] 丁蕾.基于高分辨率遥感影像的设施农用地识别技术研究[D].连云港:江苏海洋大学,2022.
- [34] 兰立波,王志冉,张栋,等.新型柔性墙体日光温室冬季温度场分布规律研究[J].*中国农业文摘·农业工程*, 2023, 35(6):36-41.
- [35] 周长吉,童根树,闫俊月,等.农业温室结构设计标准:GB/T 51424—2022[S].北京:中国计划出版社,2022.
- [36] 韩太利,魏家鹏.寿光新型日光温室的结构特点与推广应用[J].*中国蔬菜*, 2010(13):7-9.