

引用格式:何昭欣,张淼,吴炳方,等. Google Earth Engine 支持下的江苏省夏收作物遥感提取[J].地球信息科学学报,2019,21(5):752-766. [He Z X, Zhang M, Wu B F, et al. Extraction of summer crop in Jiangsu based on Google Earth Engine[J]. Journal of Geo-information Science, 2019,21(5):752-766.] DOI:10.12082/dqxxkx.2019.180420

Google Earth Engine 支持下的江苏省夏收作物遥感提取

何昭欣^{1,2}, 张 淼¹, 吴炳方^{1,2*}, 邢 强¹

1. 中国科学院遥感与数字地球研究所 遥感科学国家重点实验室, 北京 100101; 2. 中国科学院大学, 北京 100049

Extraction of Summer Crop in Jiangsu based on Google Earth Engine

HE Zhaoxin^{1,2}, ZHANG Miao¹, WU Bingfang^{1,2*}, XING Qiang¹

1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Jiangsu province, with 13 municipalities and located in the east of China, is an important part of the Yangtze river delta economy belt. The temperature is appropriate and the rainfall is moderate. Jiangsu province enjoys a moderate climate, which is suitable for the agricultural development. Winter wheat is distributed throughout the whole province, whereas the planting structure of winter rapeseed is complex and mainly scattered in Southern Jiangsu. As reported by the State Statistics Bureau, the total planting area of winter wheat and winter rapeseed in Jiangsu ranked the fifth and seventh in China, respectively, during the last 10 years. Fast obtaining the precise planting area of these two crops in Jiangsu is crucial for the agricultural development. Remote sensing classification based on local host can obtain spatial distribution of crops with high accuracy, but is time-consuming. With the development of geographical big data, cloud platform, and cloud computation, the Google Earth Engine (GEE), a global scale geospatial analysis platform based on the cloud platform, has brought new opportunities for rapid remote sensing classification. Based on the GEE cloud platform, a time-saving method of obtaining the spatial distribution of winter wheat and winter rapeseed by use of sentinel-2 data in Jiangsu was proposed. First, 119 sentinel-2 images without cloud were obtained using the GEE in Jiangsu. The time interval was set from March 1 to June 1, 2017, and the space area was Jiangsu province. Based on the spatio-temporal information, the 119 remote sensing images were mosaicked and clipped. Secondly, remote sensing indices, texture, and terrain features were calculated respectively, and the original features were extracted. The original feature space was optimized by an algorithm named Separability and Thresholds (SEaTH algorithm). Finally, four classifiers including naive Bayes, support vector machine, classification regression tree, and random forest were tested and evaluated by the average assessment accuracy. The spatial distribution

收稿日期:2018-08-29;修回日期:2019-01-11.

基金项目:国家重点研发计划项目(2016YFD0300608);中国科学院科技服务网络计划(STS计划)项目(KFJ-STZ-ZDTP-009);国家自然科学基金项目(41561144013、41861144019、41701496)。[**Foundation items:** National Key R&D Program of China, No.2016YFD0300608; Science and Technology Service Network Initiative (STS), No.KFJ-STZ-ZDTP-009; National Natural Science Foundation of China, No. 41561144013, 41861144019, 41701496.]

作者简介:何昭欣(1991-),男,山东莒县人,硕士,主要从事农作物与果园遥感分类研究。E-mail: hezx@radi.ac.cn

*通讯作者:吴炳方(1962-),男,江西玉山人,博士,研究员,主要研究领域为农业遥感与粮食安全、水资源遥感与耗水管理、生态遥感等。E-mail: wubf@radi.ac.cn

information of winter wheat and winter rapeseed were obtained quickly. The following conclusions are drawn: (1) the GEE can quickly complete pre-processing of cloud-masking, image-mosaicking, image-clipping, and feature extraction, which is superior to the local processing. (2) The distance values of J-M that are higher than 1 and rank top two highest can reduce the number of features from 28 to 11 and effectively compress the original feature space. (3) With the combined training of spectral, texture and terrain features, the average assessment accuracy of naive Bayes, support vector machine, classification regression tree, and random forest was 61%, 87%, 89% and 92%, respectively.

Key words: GEE cloud platform; Sentinel-2; feature optimization; crop identification; Jiangsu Province

***Corresponding author:** WU Bingfang, E-mail: wubf@radi.ac.cn

摘要:江苏省是农作物种植大省,国家统计局统计数据显示,江苏省近10年冬小麦、冬油菜的总播种面积分列全国第五、第七,快速准确地获取冬小麦和冬油菜的空间分布对于该省的农业发展具有重要意义。基于单机的传统遥感分类能够准确获取农作物的空间分布信息,但是耗时较长。随着地理大数据与云平台、云计算的发展,Google Earth Engine(GEE)作为一个基于云平台的全球尺度地理空间分析平台,为快速遥感分类带来了新的机遇。本文基于GEE,使用Sentinel-2数据快速提取了江苏省2017年冬小麦与冬油菜的空间分布。首先,利用GEE获得覆盖江苏省119景无云质优的Sentinel-2影像;其次,在此基础上分别计算了遥感指数、纹理特征、地形特征,并完成原始特征的构建与优化;最后,分别试验了朴素贝叶斯、支持向量机、分类回归树和随机森林4种分类器,比较了各分类器的分类精度,并提取了冬小麦与冬油菜的空间分布信息。得出以下结论:①GEE能够快速完成覆盖江苏省影像数据的去云、镶嵌、裁剪及特征构建等预处理,较本地处理具有明显优势;②J-M距离值位于前两位且大于1将特征数量从28个压缩到11个,有效压缩了原始特征空间;③光谱+纹理+地形特征组合训练,朴素贝叶斯、支持向量机、分类回归树、随机森林的平均验证精度分别为61%、87%、89%、92%。

关键词:GEE云平台;Sentinel-2;特征优选;作物识别;江苏省

1 引言

江苏省是我国冬小麦和冬油菜种植大省,据国家统计局最近10年统计,江苏省冬小麦播种面积在全国各省份中居第五,冬油菜播种面积居第七^[1]。利用遥感技术快速提取全省冬小麦与冬油菜,具有重要的现实意义。

由于受到时间、空间及影像质量等因素的影响,应用于某一研究区的遥感数据量受到限制,仅仅依靠少量的数据不能满足大面积农作物调查^[2-5]。各种遥感分类器,诸如最大似然、支持向量机、随机森林等已被实践证明,能够有效地提取某一农作物或同时提取多种农作物,但是大多都是基于单机分类,分类速度较慢^[6-8]。

Google Earth Engine(GEE)是一个基于云的行星尺度地理空间分析平台,它将谷歌最先进的云计算和存储能力用于处理各种热点社会问题,包括森林砍伐、干旱、灾害、疾病、粮食安全、水资源管理、气候监测和环境保护^[9]。作为一个集成世界各地数百万台服务器的云平台,它不仅为传统的遥感科学工作者提供支持,而且还为更广泛的公众用户提供支持。例如,在遥感分类方面,用户利用GEE可以

方便快速地提取农作物或其他地类^[10-14]。

本研究从遥感数据的使用到分类的实现,全部基于GEE云平台,主要工作包括:①遥感影像数据的使用。根据江苏省冬小麦及冬油菜的生长期、影像质量、研究区范围,利用GEE挑选了共119景Sentinel-2影像数据,并快速完成影像的镶嵌与裁剪;②特征构建。在原始Sentinel-2影像数据所有的波段基础上,通过GEE及遥感指数公式,编码实现NDVI、EVI、NDBI、NDWI、LSWI遥感指数的构建,通过GEE函数库自带的纹理特征计算函数,计算了Sentinel-2影像数据的纹理,完成纹理特征的构建,通过GEE自带的地形函数与DEM数据,计算了海拔高度、坡度、坡向及山体阴影,完成地形特征的构建;③特征优化。通过设定J-M距离大于1且位于前两位,完成类-类之间的最优特征筛选;④分类器试验。分别试验朴素贝叶斯、随机森林、支持向量机和分类回归树的分类性能,得到江苏省2017年冬小麦及冬油菜的空间分布。本论文通过GEE云平台快速完成遥感数据获取与预处理、分类器试验、作物空间分布信息提取工作,解决基于单机的传统遥感分类中,遥感数据获取困难、预处理及分类耗时的问题。

2 研究区概况与数据源

2.1 研究区概况

江苏省地处中国大陆东部,116°18'E-121°57'E, 30°45'N-35°20'N,陆域内长江、淮河经过,东邻黄海,与上海市、浙江省、安徽省和山东省接壤,是长江三角洲经济带的重要组成部分。

江苏省基本上由长江和淮河下游的大片冲积平原组成。全省的地势总体上相当低平,平原面积70 000 km²左右,约占全省面积的69%;丘陵和低山零星分布于江苏省的西南部,约占全省面积的14.3%。地处亚热带向暖温带过渡区,气候温和,雨量适中,四季气候分明。以淮河、苏北灌溉总渠一线为界,以北属暖温带湿润、半湿润季风气候,以南属亚热带湿润季风气候。全省各地平均气温13~16℃,最冷月为1月,平均气温-1.0~3.3℃,最热月为7月,平均气温26.0~28.8℃。

江苏省是冬小麦和冬油菜种植大省,如图1所示,冬小麦一般在10月底到11月初播种,至次年6月中上旬收获,种植区域在全省都有分布,主要集中在长江以北的地区,其中南通市、盐城市及泰州市分布较多。冬油菜主要在水稻收获之后在水稻田基础上种植,与冬小麦物候相似。冬油菜种植结构复杂,多以零星种植为主,主要分布在苏南地区。冬油菜是江苏省重要的油料作物,是仅次于小麦和水稻的第三大农作物。

2.2 数据来源与预处理

2.2.1 Sentinel-2影像数据

哨兵2A(Sentinel-2A)卫星是“全球环境与安全监测”计划的第二颗星,于2015年6月23日发射。该卫星携带一枚多光谱成像仪,覆盖13个光谱波段,幅宽290 km,空间分辨率10 m,重访周期5 d。在光学数据中,哨兵2A数据是唯一一个在红边范



图1 江苏省冬小麦、冬油菜物候历

Fig. 1 The phenology calendar of winter wheat and winter rape in Jiangsu province

围含有3个波段的数据。本研究使用的是Sentinel-2数据的Level-1C产品,该级别产品已经过数字高程模型校正,每个像元的值由大气顶部反射率(Top of Atmosphere, TOA)计算得到,包含13个无符号整型光谱波段,此外,还有3个QA波段,其中QA60波段包含云掩膜信息,可以用来去除影像中的卷云和厚云。有关该产品的详细波段信息见表1。

本研究使用的Sentinel-2数据存储在GEE云盘中,数据的处理全部通过在线编写代码实现。时相、轨道号挑选:根据研究区范围和冬小麦、冬油菜的物候(图1),挑选关键时相、关键位置的影像。相关研究表明,在油菜开花时期,冬小麦正处于拔节期,此时两者的光谱信息差异较大,是区分油菜和冬小麦的最佳时相^[15]。为了获得无云且包含油菜开花期的影像数据,本研究在设置影像时相上扩大了时间跨度,设置时间区间为2017年3月1日至2017年

表1 Sentinel-2波段参数

Tab. 1 The waveband parameters of Sentinel-2

序号	波段名称	中心波长/nm	空间分辨率/m	属性
01	B1	443	60	气溶胶
02	B2	490	10	蓝
03	B3	560	10	绿
04	B4	665	10	红
05	B5	705	20	红边1
06	B6	740	20	红边2
07	B7	783	20	红边3
08	B8	842	10	近红外
09	B8A	865	20	红边4
10	B9	940	60	水蒸气
11	B10	1375	60	卷云
12	B11	1610	20	短波红外1
13	B12	2190	20	短波红外2
14	QA10		10	
15	QA20		20	
16	QA60		60	云掩膜

6月1日,此区间不仅包含了2种作物光谱差异最大的信息,而且保证了影像的完整性。影像去云:计算云信息波段 QA60 的 Bit10 和 Bit11 的值,设置二者的值均为0得到云掩膜,使用云掩膜去除影像中的云信息,得到无云影像。通过以上两步,得到相应时间区间内、相应研究区的覆盖影像共119景,待后续进一步处理。

2.2.2 SRTMGL1_003数据

航天飞机雷达地形任务(The Shuttle Radar Topography Mission, SRTM^[16]),生产的数字高程数据简称SRTMGL1_003,空间分辨率为30 m,由NASA JPL提供。通过GEE中的重采样函数将其转换为与Sentinel-2数据相同的空间分辨率,并按照江苏省行政边界进行裁剪。

2.2.3 样本数据

本研究分类系统包括冬小麦、冬油菜、其他植被、水体、人工地表5类,样本的获得综合使用线上选取与线下检验2种方法。在各地类样点的选取过程中,本着随机、均匀的原则,尽量做到全省范围都有采集。由于江苏省中南部地区(如泰州市)分布有较多的冬油菜,因此在该区域选取了较多的冬油菜样点。由于冬小麦在全省范围几乎都有分布,因此在全省范围内选取了冬小麦样点。其他植被样点包含森林、草地及其他作物等。水体样点包含河流、湖泊。人工地表样点包含道路、房屋等。样本选取具体流程为:①在GEE预处理生成的Sentinel-2数据(2017年3月1日至2017年6月1日合成图)上进行各地类样点的在线选取;②将样点导出为KML格式,然后导入到Google Earth平台,使用同期高清底图对样点的正确性进行检查。同时,本研究使用样点中也包含了2017年实地采集到的样点。各地类样点的选取数量如表2所示。

3 研究方法

本文通过GEE云平台,快速获取并处理了覆盖江苏省的遥感数据,在此基础上完成了光谱、纹理及地形特征的构建与优化,试验了朴素贝叶斯、支持向量机、分类回归树及随机森林算法并评价了分类精度,最终得到江苏省2017年冬小麦、冬油菜的空间分布信息,技术流程图如图2所示。

3.1 特征构建

特征变量的选择是遥感分类的重要步骤,科学

表2 样本数据

Tab. 2 Sample data

地类代码	地类名称	样本数量/个
0	冬小麦	615
1	冬油菜	546
2	其他植被	610
3	水体	653
4	人工地表	666
总计		3090

使用多种特征变量及特征变量组合,可以有效提高遥感分类的精度^[17]。本文选择光谱特征、纹理特征及地形特征,并分别介绍如何在GEE平台上实现这3种特征的构建。

3.1.1 光谱特征

通过时相、位置、去云筛选后,共得到119景影像,每景影像包含13个原始光谱波段(表1),根据遥感专题指数计算公式编写代码,通过GEE平台计算每景影像的归一化植被指数NDVI、增强型植被指数EVI、地表水分指数LSWI^[10]、归一化水体指数NDWI^[18]、归一化建筑指数NDBI^[19]。并将每种指数作为一个独立的光谱波段添加到原始影像中。各专题指数计算公式如下:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (1)$$

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1} \quad (2)$$

$$LSWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (3)$$

$$NDWI = \frac{\rho_{GREEN} - \rho_{NIR}}{\rho_{GREEN} + \rho_{NIR}} \quad (4)$$

$$NDBI = \frac{\rho_{MIR} - \rho_{NIR}}{\rho_{MIR} + \rho_{NIR}} \quad (5)$$

式中: ρ_{MIR} 、 ρ_{SWIR} 、 ρ_{NIR} 、 ρ_{RED} 、 ρ_{GREEN} 、 ρ_{BLUE} 分别代表中红外、短波红外、近红外、绿、蓝波段的反射率值。

3.1.2 纹理特征

仅仅依靠光谱特征的农作物遥感分类效果并不理想,各类之间较容易出现混分、漏分及错分,这是因为在遥感影像上存在着“同物异谱、异物同谱”的现象^[17]。

纹理是图像的一个重要属性,是图像灰度在空间上以一定形式变化而产生的图案。灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)是一种被广泛应用的纹理统计分析方法^[20-22],其简单定义为图像上具有某种空间位置关系的两个像素灰

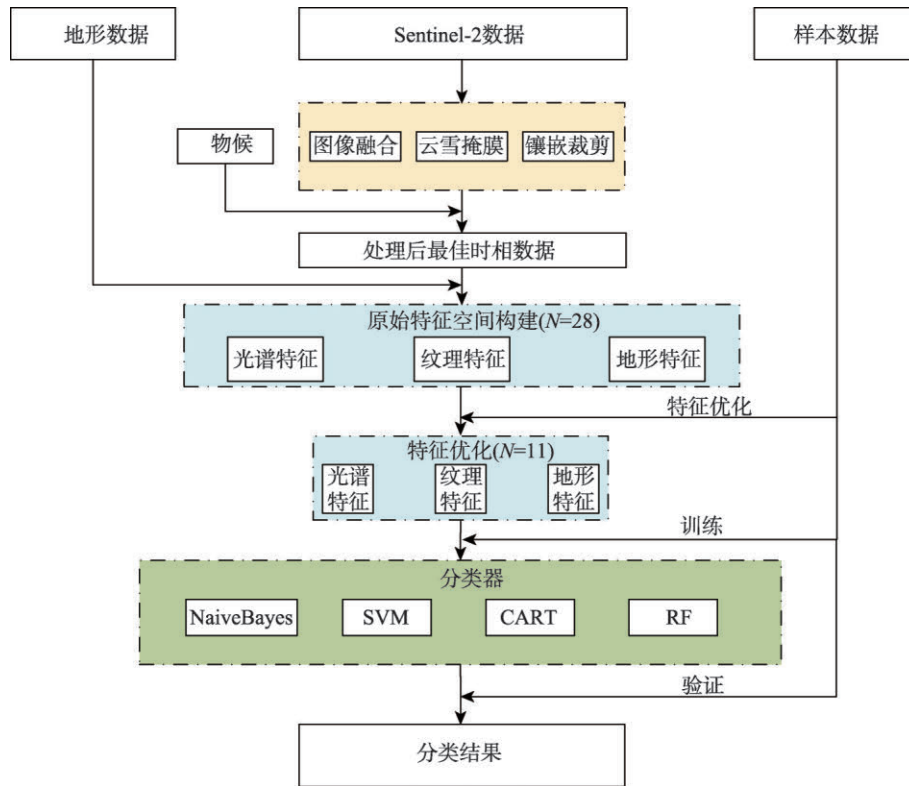


图2 夏收作物遥感提取技术流程

Fig. 2 Technical flowchart of remote sensing extraction for summer crops

度的联合概率分布。由于纹理特征是一种结构特征,使用不同波段的影像得到的纹理特征差别不大,另外由于近红外波段对植被更有效,因此本文使用 sentinel-2 影像的近红外波段 B8 计算纹理特征。GEE 提供了快速计算基于 GLCM 纹理特征的函数 `glcmTexture(size, kernel, average)`,用时仅需大约 10 s 便可以计算出角二阶矩 (B8_asm)、对比度 (B8_contrast)、相关性 (B8_corr)、方差 (B8_var)、逆差矩 (B8_idm)、和平均 (B8_savg)、和方差 (B8_svar)、和熵 (B8_sent)、熵 (B8_ent)、差方差 (B8_dvar)、差熵 (B8_dent)、相关信息测度 1 (B8_imcorr1)、相关信息测度 2 (B8_imcorr2)、最大相关系数 (B8_maxcorr)、差异性 (B8_diss)、惯性矩 (B8_inertia)、聚类熵 (B8_shade)、聚类突 (B8_prom)^[23-24]。这 18 个纹理特征全部用来训练分类器必然会产生冗余,因此原始纹理特征空间构建共选取了最常用的 6 个特征: B8_asm、B8_contrast、B8_corr、B8_var、B8_idm、B8_ent^[25-26]。

3.1.3 地形特征

遥感识别农作物,加入与地形相关的参数,能够提高分类精度。除了上述用到的纹理特征以外,

本研究使用 SRTMGL1_003 数据构建海拔高度 (elevation)、坡度 (slope)、坡向 (aspect) 及山体阴影 (hillshade) 特征。通过函数 `ee.Terrain.products(input)` 计算得到,作为 4 个独立的波段参与到原始特征的构建。

3.2 特征优化方法

可用于遥感图像分类的特征数十种甚至数百种,如果将如此多的特征全部输送到分类器去训练,不仅使运算变得复杂,处理速度大大下降,而且在有限样本的情况下,过多的特征可能会导致分类精度降低,这就是所谓的“维数灾难”^[27-28]。

通过原始特征构建,共得到光谱特征 18 个,纹理特征 6 个,地形特征 4 个,总共 28 个原始特征。本文使用的特征优化方法为 SEaTH 算法 (Separability and Thresholds), SEaTH 算法的基本思想是:在特征符合正态分布的前提下 (如果对象特征不符合正态分布规律,说明该特征分离性差,不考虑用于分类),利用 Jeffries-Matusita 距离 (简称 J-M 距离,本文用 J 表示) 来衡量类间可分性。对于类 C_1 和 C_2 类来说,其 J 距离表示为:

$$J = 2(1 - e^{-B}) \quad (6)$$

$$B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2} \ln \left[\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2} \right] \quad (7)$$

式中: B 表示巴氏距离, m_i 和 σ_i^2 , $i=1, 2$, 分别代表类 C_1 和类 C_2 某个特征的均值和方差。 J 的取值范围是 $[0, 2]$, 0 表示 2 个类别在某一特征上几乎完全混淆, 2 表示 2 个类别在某一特征上能够完全分开。然而, $J=2$ 的情况在实际操作中是不存在的, 因为各类之间总会存在一些交叠。考虑到分类模型的可移植性, 通常只保留 J 值靠前的 2 位特征参与训练与分类^[29-30]。本文的研究在此基础上提出 J 值位于前两位且 J 值大于 1 的特征才选用, 这样做主要基于二点, 一方面进一步压缩了特征数量, 另一方面 J 值大于 1 (或者说 J 值越接近 2) 的特征能够使类别之间的可分性最大。

3.3 分类方法

本研究采用 4 种不同的分类器进行训练和分类, 这 4 种分类器分别是朴素贝叶斯 (Naive Bayes)、分类回归树 (Classification and Regression Tree)、支持向量机 (Support Vector Machine) 和随机森林 (Random Forest)。

3.3.1 朴素贝叶斯 (NB)

朴素贝叶斯分类器的分类原理是通过某对象的先验概率, 利用贝叶斯公式计算出其后验概率, 选择具有最大后验概率的类作为该对象所属的类^[31-32]。

贝叶斯分类是一种简单有效的遥感影像分类算法, 它只需要少量的训练数据就可完成学习, 在分类中并不是把某一个实例绝对地归并为某一类, 而是通过计算该实例属于某一类的概率, 具有最大概率的类就是该实例所属的类。

贝叶斯分类不仅能够使测试样本的所有属性参与到分类中, 而且对样本的属性分布适用范围较广, 样本属性可以是离散的、连续的或是混合的。

3.3.2 分类回归树 (CART)

CART 决策树算法^[33]由 Breiman 提出, 该算法既可以用于分类^[34-35], 又可用于连续变量的预测, 它的基本原理是: 将训练样本划分为测试变量 (特征向量) 和目标变量 (实际地物类型), 通过对这两变量的循环分析形成二叉决策树。CART 算法采用经济学中的基尼系数 (Gini Index) 作为选择最佳测试变量和分割阈值的准则。基尼系数的定义如下:

$$GI = 1 - \sum_{j=1}^J p^2(j/h) \quad (8)$$

$$p(j/h) = \frac{n_j(h)}{n(h)} \quad (9)$$

$$\sum_{j=1}^J p(j/h) = 1 \quad (10)$$

式中: $p(j/h)$ 是从训练样本集中随机抽取一个样本, 当它的某一测试变量值为 h 时属于第 j 类的概率; $n_j(h)$ 为训练样本集中测试变量值为 h 时属于第 j 类的样本个数; $n(h)$ 为训练样本集中测试变量值为 h 的样本个数; j 为类别个数。

完整的决策树生成后并不是直接就可以用于分类, 因为它对训练样本特征的描述过于精确, 很大程度上失去了一般性, 造成“过度拟合”, 因此需要对决策树进行修剪。CART 算法采用交叉验证法进行修剪, 最终分析结果是一棵兼顾复杂度和错误率的最优二叉树。

3.3.3 支持向量机 (SVM)

支持向量机是一种重要的统计学习算法, 是基于结构风险最小化、优化和核函数的线性分类器。SVM 是非参数方法, 数据即使不满足标准概率密度分布, 该方法一样可以工作。SVM 样本分类的基本原理主要包括: ① 对于非线性可分的数据, SVM 通过算法将该数据映射到高维空间, 实现数据的线性可分; ② 在样本特征空间中通过核函数引导, 寻找分离超平面, 最大化样本中不同类别的距离。SVM 分类器具有很高的泛化能力, 且仅仅需要数量不多的样本数据, 特别适用于遥感图像的分类^[36-37]。

3.3.4 随机森林 (RF)

随机森林是分类树的集合^[38]。构建每棵分类树是随机森林算法的关键, 主要包括两个随机选择的步骤: 采用 bootstrap 抽样策略从原始数据集中创建约 2/3 的训练样本, 并为每个训练样本分别生成决策树, 剩余约 1/3 的训练样本作为 Out-of-bag 数据用于内部交叉检验, 评估随机森林的分类精度; 使用 Gini 系数确定决策树中每个节点的分裂条件, 集合每棵决策树, 完成随机森林的构建。

随机森林算法既能处理离散型数据, 也能处理连续型数据, 数据集无需规范化。两个随机性的引入, 使得随机森林不容易出现决策树的过拟合, 同时也省去了决策树的修剪, 计算负担较小。RF 的优点使得它能够处理较为复杂遥感图像分类, 适用于多类别与多特征^[39-41]。

3.4 精度评价方法

混淆矩阵(confusion matrix)也称误差矩阵,是表示精度评价的一种标准格式,用 n 行 n 列的矩阵形式来表示。具体评价指标有总体精度(Overall Accuracy, OA)、制图精度(Producer's Accuracy, PA)、用户精度(User's Accuracy, UA)及 Kappa 系数(Kappa Coefficient, KC)等,这些精度指标从不同的侧面反映了图像分类的精度^[42-43]。

4 结果与分析

4.1 特征优化结果

将 5 类地物进行两两组合,共有 10 种组合情况。在 18 个光谱特征、6 个纹理特征、4 个地形特征

中,计算每一种组合、每一种特征的 J-M 距离(表 3),以 J-M 距离大于 1 且位于前两位的特征为选择原则,最终得到最优特征 11 个(表 4)。

4.2 分类特征与分类器组合精度

本文使用朴素贝叶斯、支持向量机、分类回归树及随机森林 4 种分类器,在 GEE 中各分类器参数设置如表 5 所示。将 4 种分类器与 3 种特征组合,在江苏省夏收作物分类中进行了试验,验证精度采取交叉验证策略(表 6),具体流程为:① 在 3090 个样本点中独立生成 10 次随机数,每次生成的随机数不重复,保证样本选取的随机性;② 将所有样本进行划分,约 70% 的样本用来训练,剩余约 30% 的样本用来验证,分别得到 10 次验证精度;③ 对 10 次验证精

表 3 J-M 距离
Tab. 3 J-m distance

	0-1	0-2	0-3	0-4	1-2	1-3	1-4	2-3	2-4	3-4
B1	0.2680	0.1899	0.3451	0.8664	0.3541	0.2860	0.9974	0.0699	0.4645	0.5532
B10	0.1086	0.0510	0.2655	0.4685	0.2786	0.2367	0.7439	0.4025	0.3023	0.6612
B11	0.0345	0.2059	1.7859	0.8597	0.3667	1.7725	0.9742	1.7906	0.5136	1.5788
B12	0.0986	0.1134	0.9006	1.0067	0.3045	1.1916	1.2127	1.2405	0.8051	1.4930
B2	0.4440	0.1323	0.4348	1.0042	0.4218	0.4415	1.1720	0.1188	0.7053	0.6220
B3	0.5995	0.1026	0.2994	0.8887	0.2938	0.1648	0.7916	0.0623	0.6161	0.4861
B4	0.1689	0.0809	0.1894	1.1174	0.0806	0.0721	1.0340	0.0265	0.8892	0.8105
B5	0.8843	0.1368	0.0526	0.8874	0.4293	1.1382	0.8545	0.2567	0.5771	0.9491
B6	0.3152	0.2685	1.9948	0.4921	0.5103	1.9985	0.6967	1.6821	0.0608	1.3003
B7	0.0068	0.4568	1.9962	0.8542	0.4653	1.9980	0.8656	1.6988	0.1047	1.3183
B8	0.0303	0.3785	1.9981	0.8838	0.5434	1.9999	1.0667	1.7931	0.1511	1.4576
B8A	0.0126	0.4349	1.9984	0.9105	0.4010	1.9989	0.8751	1.8037	0.1397	1.4569
B9	0.4272	0.6179	1.9324	0.6222	0.0246	1.5326	0.0662	1.3821	0.0265	1.1396
EVI	0.0890	0.4287	1.9808	1.9302	0.4344	1.9970	1.9851	1.8262	1.5374	0.8268
LSWI	0.0513	0.4662	0.0249	1.7987	0.7753	0.1316	1.9234	0.3106	1.2424	1.6268
NDBI	0.0513	0.4662	0.0249	1.7987	0.7753	0.1316	1.9234	0.3106	1.2424	1.6268
NDVI	0.1701	0.2972	1.9492	1.9432	0.3385	1.9726	1.9892	1.7693	1.6042	0.8157
NDWI	0.3376	0.4312	1.9838	1.6811	0.4369	1.9919	1.7671	1.9134	1.0092	1.4172
B8_asm	0.0750	0.0302	1.4439	0.0103	0.1730	1.5201	0.1250	1.3745	0.0055	1.4037
B8_contrast	0.8985	0.3195	0.2373	0.5606	0.3501	0.4561	0.1051	0.0777	0.1247	0.1443
B8_corr	0.1056	0.0312	0.1653	0.0426	0.1877	0.3961	0.2288	0.0639	0.0040	0.0458
B8_ent	0.0552	0.0296	1.1539	0.0099	0.1488	1.2592	0.1035	1.0524	0.0054	1.0964
B8_idm	0.1398	0.0183	1.0954	0.0127	0.1780	1.2619	0.1310	1.0107	0.0045	1.0511
B8_var	0.8535	0.3683	0.3689	0.6562	0.2288	0.2918	0.0621	0.0559	0.1381	0.1123
aspect	0.0047	0.0084	0.1336	0.0136	0.0142	0.1159	0.0316	0.2022	0.0317	0.1383
elevation	0.1988	0.7377	0.0249	0.0298	1.0175	0.2476	0.3239	0.6190	0.5991	0.0141
hillshade	0.0098	0.4100	0.0847	0.1292	0.4665	0.1360	0.1714	0.1881	0.1197	0.0228
slope	0.0095	0.8064	0.3598	0.1896	0.8690	0.4255	0.2569	0.2386	0.4186	0.0585

注:0:冬小麦;1:冬油菜;2:其他植被;3:水体;4:人工地表。

表4 原始特征空间与优化
Tab. 4 Original features space and optimization

特征种类	原始特征	数量/个	优化后特征	数量/个
光谱特征	B1、B2、B3、B4、B5、B6、B7、B8、B8A、B9、B10、B11、B12、 <i>EVI</i> 、 <i>LSWI</i> 、 <i>NDBI</i> 、 <i>NDVI</i> 、 <i>NDWI</i>	18	B8、B8A、 <i>EVI</i> 、 <i>LSWI</i> 、 <i>NDBI</i> 、 <i>NDVI</i> 、 <i>NDWI</i>	7
纹理特征	<i>B8_asm</i> 、 <i>B8_contrast</i> 、 <i>B8_corr</i> 、 <i>B8_var</i> 、 <i>B8_idm</i> 、 <i>B8_ent</i>	6	<i>B8_asm</i> 、 <i>B8_idm</i> 、 <i>B8_ent</i>	3
地形特征	<i>elevation</i> 、 <i>slope</i> 、 <i>aspect</i> 、 <i>hillshade</i>	4	<i>elevation</i>	1
总计		28		11

表5 分类器参数
Tab. 5 Parameters of classifiers

分类器	参数
朴素贝叶斯	Lambda:0.000 000 1
支持向量机	kernelType:RBF; gamma:0.5; cost:50
分类回归树	crossvalidationFactor:10; maxDepth:10
随机森林	NumberOfTrees:200; variablesPerSplit:0

度求平均,作为最终的平均验证精度(表6、图3)。

4.3 冬小麦、冬油菜遥感提取与精度验证

图4为遥感提取的2017年江苏省冬小麦、冬油菜、水体、其他(其他植被、人工地表),由于本论文的主要目的是提取冬小麦和冬油菜,为了突出这两

表6 不同特征、不同分类器组合的平均验证精度分布
Tab. 6 Average verification precisions of different combinations of features and classifiers

分类器		特征						
		光谱	纹理	地形	光谱+纹理	光谱+地形	纹理+地形	光谱+纹理+地形
朴素贝叶斯	0	0.49	0.37	0.21	0.49	0.59	0.29	0.59
	1	0.50	0.21	0.21	0.50	0.60	0.21	0.60
	2	0.51	0.21	0.21	0.51	0.61	0.21	0.65
	3	0.51	0.20	0.20	0.51	0.60	0.20	0.60
	4	0.49	0.21	0.21	0.49	0.60	0.21	0.60
	5	0.49	0.21	0.21	0.49	0.59	0.21	0.65
	6	0.51	0.21	0.21	0.51	0.60	0.21	0.61
	7	0.49	0.21	0.21	0.49	0.58	0.21	0.60
	8	0.50	0.19	0.19	0.50	0.61	0.19	0.61
	9	0.50	0.20	0.20	0.50	0.58	0.20	0.58
支持向量机	均值	0.50	0.22	0.21	0.51	0.60	0.21	0.61
	0	0.85	0.23	0.38	0.86	0.85	0.38	0.86
	1	0.87	0.21	0.38	0.87	0.84	0.37	0.87
	2	0.86	0.21	0.40	0.88	0.88	0.41	0.88
	3	0.85	0.20	0.42	0.85	0.86	0.41	0.86
	4	0.84	0.21	0.41	0.86	0.85	0.41	0.86
	5	0.85	0.21	0.40	0.87	0.86	0.38	0.89
	6	0.86	0.21	0.40	0.88	0.85	0.38	0.89
	7	0.85	0.21	0.38	0.84	0.84	0.40	0.86
	8	0.85	0.19	0.38	0.85	0.87	0.37	0.84
分类回归树	9	0.86	0.20	0.40	0.88	0.85	0.40	0.88
	均值	0.85	0.21	0.39	0.86	0.86	0.39	0.87
	0	0.82	0.39	0.38	0.82	0.89	0.47	0.89
	1	0.84	0.40	0.37	0.84	0.86	0.49	0.88
	2	0.86	0.41	0.39	0.85	0.90	0.48	0.91
	3	0.82	0.40	0.40	0.82	0.88	0.51	0.91
	4	0.84	0.41	0.39	0.83	0.88	0.51	0.89
	5	0.84	0.39	0.40	0.86	0.88	0.48	0.90

续表6

分类器		特征						
		光谱	纹理	地形	光谱+纹理	光谱+地形	纹理+地形	光谱+纹理+地形
分类回归树	6	0.84	0.39	0.38	0.84	0.86	0.49	0.86
	7	0.83	0.39	0.37	0.84	0.88	0.49	0.89
	8	0.82	0.37	0.38	0.82	0.87	0.50	0.87
	9	0.84	0.41	0.40	0.85	0.87	0.50	0.88
	均值	0.83	0.40	0.39	0.84	0.88	0.49	0.89
随机森林	0	0.87	0.39	0.37	0.87	0.90	0.51	0.90
	1	0.88	0.40	0.38	0.89	0.91	0.55	0.92
	2	0.89	0.39	0.40	0.89	0.93	0.54	0.93
	3	0.88	0.39	0.39	0.89	0.90	0.52	0.91
	4	0.88	0.41	0.39	0.88	0.89	0.54	0.90
	5	0.90	0.43	0.38	0.90	0.91	0.53	0.92
	6	0.87	0.41	0.37	0.88	0.90	0.52	0.91
	7	0.88	0.39	0.46	0.88	0.91	0.55	0.92
	8	0.86	0.41	0.37	0.88	0.89	0.54	0.92
	9	0.88	0.37	0.38	0.88	0.90	0.52	0.91
	均值	0.88	0.40	0.39	0.89	0.90	0.53	0.92

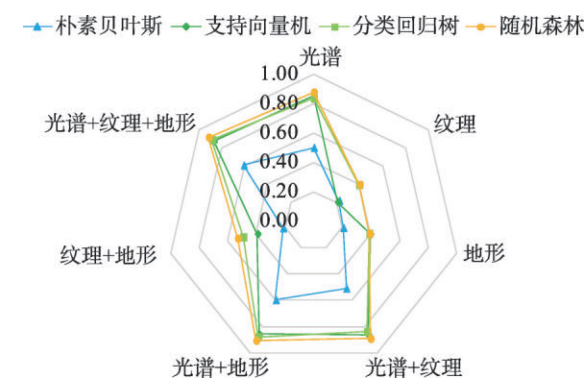


图3 不同特征、分类器组合的平均验证精度雷达分布
Fig. 3 Average verification precision radar distribution map with different combinations of features and classifiers

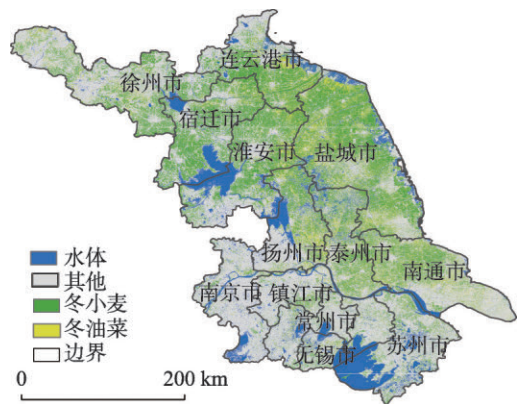


图4 2017年江苏省冬小麦和冬油菜空间分布
Fig. 4 Spatial distribution of winter wheat and winter rape in Jiangsu province in 2017

类,故只保留了水体,其他植被和人工地表显示为其他。可以看出江苏省的冬小麦和冬油菜主要分布在长江以北的城市,包括徐州市、连云港市、宿迁市、淮安市、盐城市、扬州市、泰州市和南通市,位于长江以南的南京市、镇江市、常州市、无锡市和苏州市零星分布有少量的冬小麦和冬油菜。

使用光谱+纹理+地形这3种特征组合,以随机森林算法为例,将第三次的精度验证以混淆矩阵的形式显示(表7),冬小麦的制图精度为83%,冬小麦真实样本共187个,其中155个被正确分类,32个被错分成其他类别。冬油菜的制图精度为94%,冬油菜真实样本共171个,其中160个被正确分类,11个被错分成其他类别。总体验证精度为93%,kappa系数为0.92。

5 讨论

5.1 GEE 云计算

GEE 不仅存储了大量的地理空间数据,而且开发集成了许多处理这些数据的程序。GEE 凭借其强大的云计算平台,能够高效处理地理空间数据。

以往分类所用遥感数据量有限^[44-47],想要获得同时满足空间及时间要求的影像往往不能做到。如果研究区的范围足够大,单景影像不能完全覆盖,就需要进行镶嵌,而影像镶嵌又可能会造成色彩分布不

表7 随机森林算法精度评价混淆矩阵

Tab. 7 Precision evaluation confusion matrix of Random Forest algorithm

混淆矩阵		真实类别					总计	用户精度
		冬小麦	冬油菜	其他植被	水体	人工地表		
预测类别	冬小麦	155	7	2	0	1	165	0.94
	冬油菜	10	160	2	0	0	172	0.93
	其他植被	18	3	168	0	1	190	0.88
	水体	2	0	0	203	4	209	0.97
	人工地表	2	1	8	1	192	204	0.94
	总计	187	171	180	204	198	940	
	制图精度	0.83	0.94	0.93	1.00	0.97		
	总体精度	0.93						
	Kappa系数	0.92						

均匀的问题。本研究凭借 GEE 存储大量遥感数据的优势,聚合了 2017 年 3 月 1 日至 6 月 1 日覆盖江苏省的 119 景 Sentinel-2 数据影像,数据量大。通过影像头文件自带的云信息,挑选无云清晰的高质量影像,同时实现了影像镶嵌、裁剪等一系列预处理,有效避免了影像镶嵌带来的色差问题,得到全省 Sentinel-2 影像(图 5),保证了分类数据源的纯洁度与质量。

5.2 分类结果的不确定性分析

本研究使用的是 Sentinel-2 影像,影像空间分辨率 10 m。江苏省地处江淮平原、江湖较多,在省北部和西南部分布着低山丘陵,特殊的地形造成江苏省的耕地地块比较破碎,有可能造成影像上 1 个像元内覆盖有多种农作物,加大了分类难度,造成分类误差。

(1)在地类采样过程中,由于人为因素的存在,

加之受自然环境的影响,不可能保证采到的样本完全准确,而且样本如果分布不均匀,对分类器也会造成一定的干扰,影响训练精度。

(2)卫星在获取地面影像时,也受到诸多因素的影响,既有来自本身的影响,例如传感器参数、卫星运行轨迹、影像处理误差等,也有来自外部因素的影响,例如气象条件。以上内外因都会对最终的分

类影像造成不可避免的几何及辐射误差。

(3)“同物异谱”和“异物同谱”现象的存在是影响分类精度最重要的原因。如图 6 所示,有些冬小麦被错分成了冬油菜,而有些冬油菜被错分成了冬小麦,另外其他植被也容易和冬小麦及冬油菜混分,水体和人工地表的混分最不明显。

通过计算每种地类的特征均值,得到每种地类的特征曲线。冬小麦、冬油菜的光谱特征较为接近,导致二者容易混分。冬小麦、冬油菜与其他类

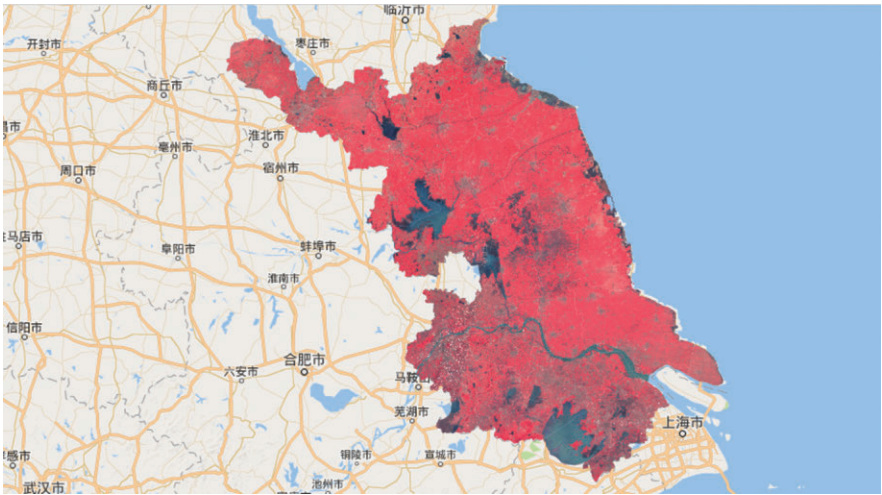


图5 江苏GEE影像处理

Fig. 5 The GEE image processing for Jiangsu

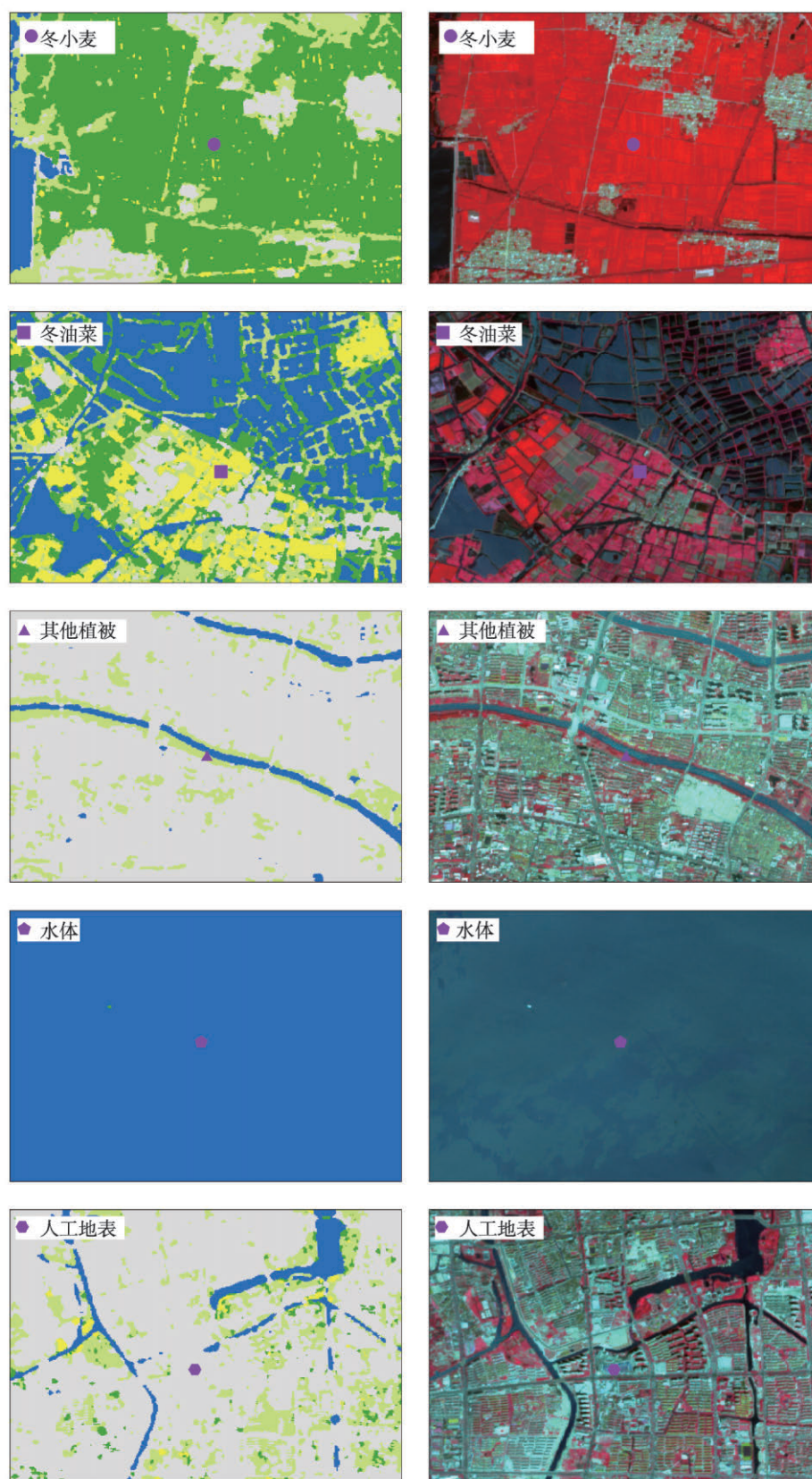


图6 Sentinel-2影像随机森林分类结果(左图)与原始影像(右图)的对比

Fig. 6 Random forest classification results of Sentinel-2 image compared with the original image

别的特征差别较为明显,尤其是在B8、B8A、NDVI等光谱特征上的差异性较大,可以利用这些特征对地类进行区分(图7)。5种地类的角二阶矩(B8_asm)、相关性(B8_corr)、熵(B8_ent)、逆差矩(B8_idm)较为接近,这些特征的相似性使得地类区分难度上升(图8)。5种地类在海拔高度(elevation)和坡向(aspect)之间的差异较为明显,可以用这两个特征来区分不同地类(图9)。

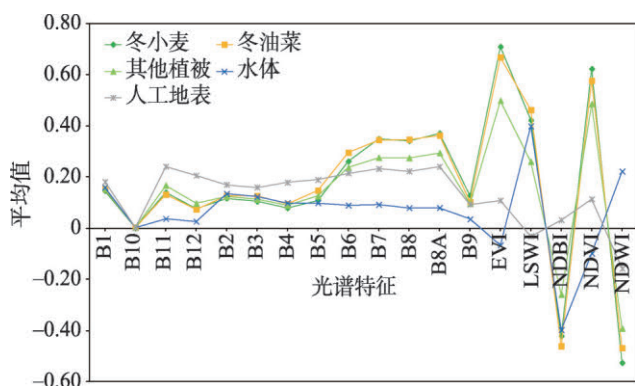


图7 不同类别的光谱特征曲线

Fig. 7 Spectral feature graph of different classes

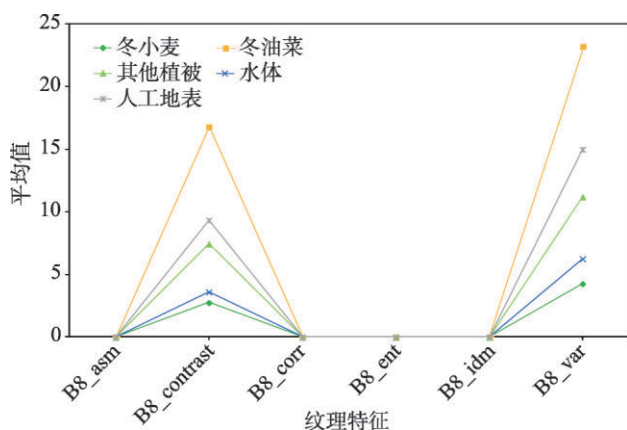


图8 不同类别的纹理特征曲线

Fig. 8 Texture feature graph of different classes

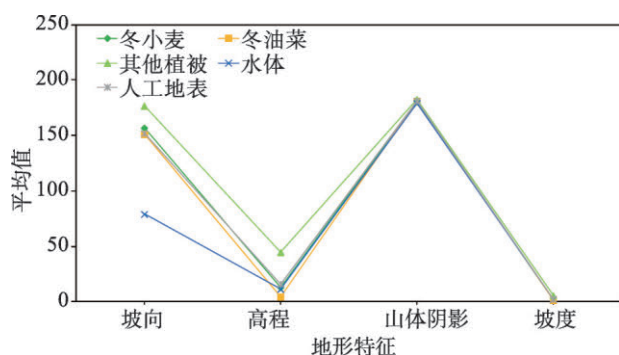


图9 不同类别的地形特征曲线

Fig. 9 Topographic feature graph of different classes

5.3 特征对精度的影响分析

本文基于GEE云平台,使用朴素贝叶斯、支持向量机、分类回归树、随机森林,及光谱特征、纹理特征、地形特征,在江苏省夏粮作物提取中进行了分类试验。总体来看,在四种分类器中,纹理、地形、纹理+地形这3种特征参与分类的平均验证精度不高,而光谱、光谱+纹理、光谱+地形、光谱+纹理+地形这四种特征参与分类的平均验证精度较高。以随机森林算法为例,纹理、地形、纹理+地形3种特征的平均验证精度分别为40%、39%、53%。光谱、光谱+纹理、光谱+地形、光谱+纹理+地形四种特征分别取得了88%、89%、90%、92%的平均验证精度(表6)。以上结果说明,光谱特征在分类中起着决定性作用,而仅仅使用单一纹理或者地形特征不能取得较高的验证精度。

另外,虽然江苏省大部分地区都是平原,但是海拔高度特征的加入使得平均验证精度得以提升。以随机森林算法为例,仅使用光谱特征,平均验证精度为88%,加入海拔高度特征后,精度提升了2%;使用光谱+纹理特征的平均验证精度为89%,加入海拔高度特征后使得精度提升了3%。从图10看出,江苏省内,冬小麦、冬油菜、其他植被、水体、人工地表的平均海拔高度不一,其他植被(森林、草地等)的海拔高度最大,这也与实际情况相符合。正是由于各地类的海拔高度不一,所以海拔高度这一地形特征的加入,有利于区分不同地类。

6 结论

本研究基于Google Earth Engine云平台,快速得到覆盖江苏省的2017年3月1日至2017年6月1日期间的Sentinel-2遥感影像并完成数据的预处理,构建了光谱、纹理、地形特征并采取SEaTH算法优化

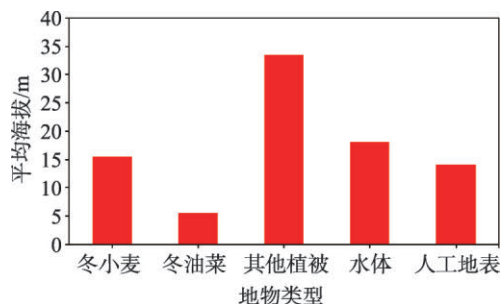


图10 不同类别的平均海拔高度

Fig. 10 Average altitude of different classes

了特征,试验了朴素贝叶斯、支持向量机、分类回归树、随机森林4种算法并比较了分类精度,快速提取了江苏省2017年冬小麦与冬油菜的空间位置分布信息。主要结论如下:

(1) GEE可以方便地访问海量遥感图像及其他数据资源,并能以其高性能的云计算能力处理这些数据。GEE能够快速完成覆盖江苏省影像数据的去云、镶嵌、裁剪及特征构建等预处理,较本地处理具有明显优势。

(2) 江苏省冬小麦和冬油菜遥感提取实验证明,J-M距离大于1且位于前两位可以有效地筛选最优特征及特征组合,将特征数量从28个压缩到11个,在降低分类器负荷的同时,提高了分类精度。

(3) 光谱特征作为分类的重要特征,对分类精度起着决定性作用,加入适当的纹理或者地形特征能提高仅使用光谱特征分类的精度,而仅使用纹理特征或者地形特征,不能取得理想的分类精度。使用光谱+纹理+地形特征的组合,朴素贝叶斯、支持向量机、分类回归树、随机森林的平均验证精度分别为61%、87%、89%、92%。

然而,基于像元进行分类将不可避免地出现了“椒盐现象”,下一步研究重点将放在面向对象分类与GEE云计算平台的结合上,通过借鉴一些优秀成熟的分割算法,实现江苏省乃至其他省份或地区的作物快速提取。

参考文献(References):

- [1] 国家统计局.分省年度数据,2008-2017[EB/OL].<http://data.stats.gov.cn/easyquery.htm?cn=E0103>. [State Statistics Bureau. Provincial annual data, 2008-2017[EB/OL]. <http://data.stats.gov.cn/easyquery.htm?cn=E0103>.]
- [2] 郭交,朱琳,靳标.基于Sentinel-1和Sentinel-2数据融合的农作物分类[J].农业机械学报,2018,49(4):192-198. [Guo J, Zhu L, Jin B. Crop classification based on data fusion of Sentinel-1 and Sentinel-2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018,49(4): 192-198.]
- [3] 史飞飞,高小红,杨灵玉,等.基于HJ-1A高光谱遥感数据的湟水流域典型农作物分类研究[J].遥感技术与应用, 2017,32(2):206-217. [Shi F F, Gao X H, Yang L Y, et al. Research on typical crop classification based on HJ-1A hyperspectral data in the Huangshui river basin[J]. Remote Sensing Technology & Application, 2017,32(2):206-217.]
- [4] 班松涛.县域农作物分类类型遥感识别与提取[D].西安:西北农林科技大学,2014. [Ban S T. Remote sensing identification and extraction of county crop classification types[D]. Xi'an: Northwest A&F University, 2014.]
- [5] 程彬.基于支持向量机的乾安县土地利用遥感分类研究[J].长春师范大学学报,2017,36(12):86-88. [Cheng B. Land use information extraction based on support vector machine using multitemporal remote sensing in Qian'an county[J]. Journal of Changchun Normal University, 2017,36(12):86-88.]
- [6] 关舒婧,韩鹏鹏,王月如,等.华南地区典型种植园地遥感分类研究[J].地球信息科学学报,2017,19(11):1538-1546. [Guan S J, Han P P, Wang Y R, et al. Study on the classification of typical plantations in south China[J]. Journal of Geo-information Science, 2017,19(11):1538-1546.]
- [7] 张静,张翔,田龙,等.西北旱区遥感影像分类的支持向量机法[J].测绘科学,2017,42(1):49-52. [Zhang J, Zhang X, Tian L, et al. The support vector machine method for RS images' classification in northwest arid area[J]. Science of Surveying and Mapping, 2017,42(1):49-52.]
- [8] 杨屹鸥,蒋平安,武红旗,等.基于高分一号与Landsat 8卫星影像的库尔勒市香梨种植面积识别研究[J].山东农业科学,2018,50(1):153-157. [Yang Y K, Jiang P A, Wu H Q, et al. Recognition research of Korla fragrant pear planting area based on satellite images of GF-1 and Landsat 8[J]. Shandong Agricultural Sciences, 2018,50(1):153-157.]
- [9] Gorelick N, Hancher M, Dixon M, et al. Google earth engine: planetary-scale geospatial analysis for everyone[J]. Remote Sensing of Environment, 2017,202:18-27.
- [10] Dong J, Xiao X, Menarguez MA, et al. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and google earth engine [J]. Remote Sensing of Environment, 2016,185:142-154.
- [11] 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. [Xu H Z Y, Liu C, Wang J B, et al. Study on extraction of citrus orchard in Gannan region based on google earth engine platform[J]. Journal of Geo-Information Science, 2018,20(3):396-404.]
- [12] Shelestov A, Lavreniuk M, Kussul N, et al. Exploring google earth engine platform for big data processing: classification of multi-temporal satellite imagery for crop mapping[J]. Frontiers in Earth Science, 2017,5(5):1-10.
- [13] 胡云锋,商令杰,王召海,等.GEE平台和CART方法的北京市土地解译[J].测绘科学,2018,43(4):87-93. [Hu Y F, Shang L J, Wang Z H, et al. Land-cover and land-use classification in Beijing based on cart and gee[J]. Science of Surveying and Mapping, 2018,43(4):87-93.]
- [14] Xiong J, Thenkabail P S, Gumma M K, et al. Automated

- cropland mapping of continental Africa using google earth engine cloud computing[J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2017,126:225-244.
- [15] 王松林.基于多源遥感的江苏省油菜及其同物候期主要作物面积提取[D].武汉:长江大学,2015. [Wang S L. Extraction of main crop areas in rapeseed and its phenological period in Jiangsu province based on multi-source remote sensing[D]. Wuhan: Yangtze University, 2015.]
- [16] Farr T G, Rosen P A, Caro E, et al. The shuttle radar topography mission[J]. *Reviews of Geophysics*, 2007,45(2):361.
- [17] 贾坤,李强子.农作物遥感分类特征变量选择研究现状与展望[J].*资源科学*,2013,35(12):2507-2516. [Jia K, Li Q Z. Review of features selection in crop classification using remote sensing data[J]. *Resources Science*, 2013,35(12):2507-2516.]
- [18] Yang X, Zhao S, Qin X, et al. Mapping of urban surface water bodies from Sentinel-2 msi imagery at 10 m resolution via ndwi-based image sharpening[J]. *Remote Sensing*, 2017,9(6):596-614.
- [19] Li K, Chen Y. A genetic algorithm-based urban cluster automatic threshold method by combining viirs dnb, ndvi, and ndbi to monitor urbanization[J]. *Remote Sensing*, 2018,10(2):277-297.
- [20] 裴欢,孙天娇,王晓妍.基于Landsat 8 OLI影像纹理特征的面向对象土地利用/覆盖分类[J].*农业工程学报*,2018,34(2):248-255. [Pei H, Sun T J, Wang X Y. Object-oriented land use/cover classification based on texture features of Landsat 8 OLI image[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2018,34(2):248-255.]
- [21] 张博.基于灰度共生矩阵的遥感影像纹理特征提取分析[J].*城市地理*,2017(22):190-200. [Zhang B. Texture feature extraction analysis of remote sensing image based on gray level co-occurrence matrix[J]. *Cultural Geography*, 2017(22):190-200.]
- [22] 林雪,彭道黎,黄国胜,等.结合多尺度纹理特征的遥感影像面向对象分类[J].*测绘工程*,2016,25(7):22-27. [Lin X, Peng D L, Huang G S, et al. Object-oriented classification with multi-scale texture feature based on remote sensing image[J]. *Engineering of Surveying and Mapping*, 2016,25(7):22-27.]
- [23] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification[J]. *IEEE Transactions on systems, man, and cybernetics*, 1973,3(6):610-621.
- [24] 方天红,陈庆虎,鄢煜尘,等.基于统计纹理特征的打印文档认证[J].*武汉大学学报(工学版)*,2016,49(1):154-160. [Fang T H, Chen Q H, Yan Y C. et al. Print document identification based on statistical texture feature[J]. *Engineering Journal of Wuhan University*, 2016,49(1):154-160.]
- [25] 冯建辉,杨玉静.基于灰度共生矩阵提取纹理特征图像的研究[J].*北京测绘*,2007,84(3):19-22. [Feng J H, Yang Y J. Study of texture images extraction based on gray level co-occurrence matrix[J]. *Beijing Surveying and Mapping*, 2007,84(3):19-22.]
- [26] 焦蓬蓬,郭依正,刘丽娟,等.灰度共生矩阵纹理特征提取的Matlab实现[J].*计算机技术与发展*,2012,22(11):169-171. [Jiao P P, Guo Y Z, Liu L J, et al. Implementation of gray level co-occurrence matrix texture feature extraction using matlab[J]. *Computer Technology and Development*, 2012,22(11):169-171.]
- [27] 张秀英,冯学智,江洪.面向对象分类的特征空间优化[J].*遥感学报*,2009,13(4):659-669. [Zhang X Y, Feng X Z, Jiang H. Feature set optimization in object-oriented methodology[J]. *Journal of Remote Sensing*, 2009,13(4):659-669.]
- [28] De Sa J P M. Pattern recognition: Concepts, methods and applications[M]. Berlin: Springer Science & Business Media, 2012.
- [29] Nussbaum S, Niemeyer I, Canty M J. Seath-a new tool for automated feature extraction in the context of object-based image analysis[C]//1st International Conference on Object-based Image Analysis, Salzburg: Austria. 2006.
- [30] 余晓敏,湛飞并,廖明生,等.利用改进SeATH算法的面向对象分类特征选择方法[J].*武汉大学学报·信息科学版*, 2012,37(8):921-924. [Yu X M, Zhan F B, Liao M S, et al. Object-oriented feature selection algorithms based on improved Seath algorithms[J]. *Geomatics and Information Science of Wuhan University*, 2012,37(8):921-924.]
- [31] Bruzzone L. An approach to feature selection and classification of remote sensing images based on the bayes rule for minimum cost[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2000,38(1):429-438.
- [32] Yang J, Ye Z, Zhang X, et al. Attribute weighted naive bayes for remote sensing image classification based on cuckoo search algorithm[C]//International Conference on Security, Pattern Analysis, and Cybernetics. IEEE, 2017:169-174.
- [33] Breiman L, Friedman J H, Olshen R, et al. Classification and regression trees[J]. *Encyclopedia of Ecology*, 1984,40(3):582-588.
- [34] Bittencourt H R, Clarke R T. Use of classification and regression trees (cart) to classify remotely-sensed digital images[C]//Geoscience and Remote Sensing Symposium, 2003. Igarss '03. Proceedings. 2003 IEEE International. IEEE, 2003,6:3751-3753.
- [35] Zhou W, Song Y Q, Pan Z K, et al. Classification of urban construction land with worldview-2 remote sensing image based on classification and regression tree algorithm

- [C]// Computational Science and Engineering (CSE) and Embedded and Ubiquitous Computing (EUC), 2017 IEEE International Conference on. IEEE, 2017,2:277-283.
- [36] Mallet C, Bretar F, Roux M, et al. Relevance assessment of full-waveform lidar data for urban area classification [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2011,66(6):S71-S84.
- [37] Gulnaz A, Sun T L, Liang Y, et al. A new technique for remote sensing image classification based on combinatorial algorithm of svm and knn[J]. *International Journal of Pattern Recognition & Artificial Intelligence*, 2018,32(7):1-23.
- [38] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1):5-32.
- [39] Feng Q, Liu J, Gong J. Uav remote sensing for urban vegetation mapping using random forest and texture analysis [J]. *Remote Sensing*, 2015,7(1):1074-1094.
- [40] 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):297-303. [Ma Y, Jiang Q G, Meng Z G, et al. Classification of land use in farming area based on random forest algorithm[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016,47(1):297-303.]
- [41] Rodriguez-galiano V F, Chica-olmo M, Abarca-hernandez F, et al. Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture[J]. *Remote Sensing of Environment*, 2012, 121(138):93-107.
- [42] 侯西勇,邸向红,侯婉,等.中国海岸带土地利用遥感制图及精度评价[J].*地球信息科学学报*,2018,20(10):1478-1488. [Hou X Y, Di X H, Hou W, et al. Accuracy evaluation of land use mapping using remote sensing techniques in coastal zone of China[J]. *Journal of Geo-information Science*, 2018,20(10):1478-1488.]
- [43] Huang D, Xu S, Sun J, et al. Accuracy assessment model for classification result of remote sensing image based on spatial sampling[J]. *Journal of Applied Remote Sensing*, 2017,11(4):1-13.
- [44] Melville B, Lucieer A, Aryal J. Object-based random forest classification of Landsat etm+ and Worldview-2 satellite imagery for mapping lowland native grassland communities in Tasmania, Australia[J]. *International Journal of Applied Earth Observation & Geoinformation*, 2018, 66:46-55.
- [45] 杨文杰.基于Landsat 8生长时序遥感信息的玉米干旱监测研究[D].石河子:石河子大学,2017. [Yang W J. Maize drought monitoring based on Landsat 8 growth time series remote sensing information[D]. Shihezi: Shihezi University, 2017.]
- [46] 陈威.基于Hadoop的K-means遥感影像分类算法的研究[D].赣州:江西理工大学,2017. [Chen W. Research on k-means remote sensing image classification algorithm based on hadoop[D]. Ganzhou: Jiangxi University of Science and Technology, 2017.]
- [47] 乌兰,乌兰吐雅,包珺玮.基于Landsat 8 OLI影像的大兴安岭西麓春小麦识别方法的比较研究[J].*北方农业学报*,2017,45(2):109-112. [Wu L, Wu L T Y, Bao J W. Study on comparison of identification methods for spring wheat at Great Khingan western slope based on Landsat 8 OLI[J]. *Journal of Northern Agriculture*, 2017,45(2):109-112.]