

刘俊伟, 陈鹏飞, 张东彦, 等. 基于时序 Sentinel-2 影像的梨树县作物种植结构[J]. 江苏农业学报, 2020, 36(6): 1428-1436.
doi: 10.3969/j.issn.1000-4440.2020.06.011

基于时序 Sentinel-2 影像的梨树县作物种植结构

刘俊伟^{1,2}, 陈鹏飞^{1,3}, 张东彦², 赵红伟⁴

(1.中国科学院地理科学与资源研究所/资源与环境信息系统国家重点实验室, 北京 100101; 2.安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心, 安徽 合肥 230601; 3.江苏省地理信息资源开发与利用协同创新中心, 江苏南京 210023; 4.中国农业科学院农业资源与农业区划研究所, 北京 100081)

摘要: Sentinel-2 影像具有空间分辨率高, 重访周期短的优势。本研究探讨了基于 Sentinel-2 数据开展梨树县作物精准分类, 进而进行种植结构分析的可行性。为此, 收集了多时相的 Sentinel-2 影像和大量地面不同作物类型样点数据。基于以上数据, 选择决策树法、最大似然法、支持向量机法等 3 种经典分类方法开展影像分类, 并对它们的结果进行对比以选择最优分类方法。然后, 利用最优分类方法获得的分类结果对梨树县作物种植结构进行分析。结果表明, 基于时序 Sentinel-2 影像, 利用作物的物候特征与其光谱特征之间的联系可实现对梨树县作物种植面积和空间分布信息的准确提取, 从而对该县种植结构进行客观评价。3 种分类方法中, 决策树法的分类精度最高, 其总体分类精度为 93.53%, Kappa 系数达到 0.890 6。

关键词: 作物种植结构; Sentinel-2; 光谱特征

中图分类号: S127 **文献标识码:** A **文章编号:** 1000-4440(2020)06-1428-09

Crop planting structure in Lishu County based on time-series Sentinel-2 images

LIU Jun-wei^{1,2}, CHEN Peng-fei^{1,3}, ZHANG Dong-yan², ZHAO Hong-wei⁴

(1. Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences/State Key Laboratory of Resources and Environmental Information System, Beijing 100101, China; 2. National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230601, China; 3. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China; 4. Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

Abstract: Sentinel-2 images have the advantages of high spatial resolution and short revisit cycle. This study investigated the feasibility of making accurate classification of crops and planting structure analysis in Lishu County based on Sentinel-2 data. Multi-temporal Sentinel-2 images and a large number of ground sample data of different crops were collected. Based on the above data, three classic classification methods including decision tree method, maximum likelihood method and support Vector machine method were selected in image classification, and their results were compared to select the

收稿日期: 2020-04-10

基金项目: 国家重点研发计划项目 (2017YFD0201501); 国家自然科学基金项目 (41871344); 中国科学院战略先导科技专项 (XDA23100100)

作者简介: 刘俊伟 (1996-), 男, 安徽合肥人, 硕士研究生, 研究方向为农业遥感。(Tel) 15979096269; (E-mail) 844385476@qq.com

通讯作者: 陈鹏飞, (Tel) 13811997067; (E-mail) pengfeichen@igsrr.ac.cn

best classification method. Then, the classification results obtained by the best method were used to analyze the crop planting structure of Lishu County. The results showed that based on time-series Sentinel-2 images, the planting area and spatial distribution information of crops in Lishu County could be extracted accurately by combining the phenological characteristics and spectral characteristics of the crops, then the planting structure of the county could be evaluated objectively. Among the three classification meth-

ods, the decision tree method got the best results, with the overall classification accuracy of 93.53% and the Kappa coefficient value of 0.890 6.

Key words: crop planting structure; Sentinel-2; spectral characteristics

作物的种植结构能够体现人们对某一地区农业生产资源的利用情况^[1]。及时、准确地掌握作物种植结构可为区域作物结构调整和优化提供有效依据^[2]。梨树县是中国重要的商品粮生产基地^[3],素有“东北粮仓”和“松辽明珠”之称,因此掌握梨树县作物的种植结构信息对农业生产管理、农业可持续发展及保障农民收益最大化等都具有重要意义。

对于农作物种植结构的研究主要有数据统计和遥感监测 2 种方法。其中,数据统计法采用区域逐层汇总统计的步骤,存在主观因素较大,容易造成统计不准确的缺陷,并且当统计对象区域范围较大时,该方法耗时耗力。遥感监测是基于卫星遥感影像,根据不同作物反射光谱信息的差异来进行分类,进而获取研究区农作物的空间分布信息。与统计方法相比,其结果更为客观,而且效率更高。目前,已有大量基于遥感技术对县级尺度作物进行分类的报道。基于多时相 Landsat8 影像,张荣群等^[4]根据曲周县主要农作物的物候特征,通过分析 *NDVI* 的时序变化,提取了该县主要农作物的空间分布及面积;刘珣等^[5]基于多时相 HJ 卫星 CCD 影像获取了嘉祥县主要秋收作物的种植结构信息。已有的类似研究大多基于 MODIS、AVHRR、Landsat、HJ 等遥感数据源,使用它们的时序数据并利用不同作物物候信息的差异展开分类。与以上提及的数据源相比,Sentinel-2 数据具有空间分辨率高,重访周期短的优势。Sentinel-2 卫星服役时间不长,其数据在土地覆盖监测、生物量估计、风险灾害预测等方面的应用正在陆续开展中^[6-9]。王大钊等^[10]为了监测鄱阳湖的洪涝灾害情况,分别利用 Sentinel-2 和 Landsat8 影像提取了湖泊的水体分布信息,结果显示使用 Sentinel-2 影像的提取精度相较于 Landsat8 高 4%,体现了 Sentinel-2 数据在地物面积提取方面的优势。刘怀鹏等^[11]基于 Sentinel-2A 数据,利用光谱指数密度分割法与光谱波段分类法成功提取了内蒙古五原县郊区的村落面积。这些研究推动了 Sentinel-2 数据的应用,但总的来说还需要开展广泛的验证,以评估其在农作物分类方面的能力和应用潜力。

基于多时相 Sentinel-2 影像,本研究以梨树县为

试验区,探讨基于 Sentinel-2 数据开展梨树县农作物分类,进行种植结构分析的可行性,以期推动高分辨 Sentinel-2 影像应用于县域尺度作物分类,提高作物分类、种植结构分析等相关研究的精度。

1 材料与方法

1.1 研究区概况

梨树县隶属于吉林省四平市(123°45′~124°53′E,43°02′~43°46′N),辖 15 个镇、6 个乡,面积为 4 209 km²。其位于长白山区向草原过渡地带^[12],地势成阶梯状分布,南部地区为低山丘陵,中部及北部地区为平原。属北温带半湿润大陆季风性气候,日照时间长,降水充沛,且雨热同季^[13]。梨树县土壤肥沃,有机质含量多,主要类型为黑土和黑钙土^[14],适宜一年一熟作物的生长,主要种植玉米、水稻、大豆、蔬菜等。研究区的地理位置如图 1 所示。

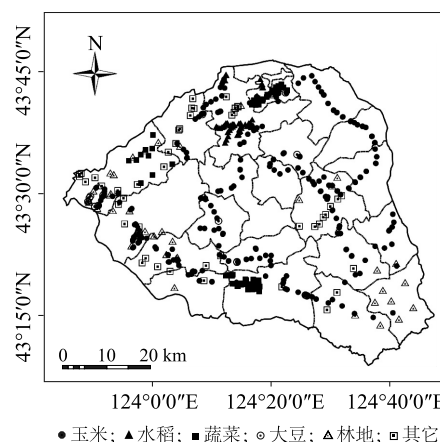


图 1 梨树县地理位置及样点分布

Fig.1 Geographical location of Lishu County and sample distribution

1.2 数据源

1.2.1 遥感数据及预处理 本研究采用的是 Sentinel-2 影像数据。它含有 2A 和 2B 两颗卫星,两星组网后重访周期为 5 d^[15]。Sentinel-2 卫星携带的多光谱成像仪(MSI)共有 13 个波段,能形成 10 m、20 m、60 m 3 种分辨率的影像,本研究在进行作物分类时,使用了 10 m 分辨率的影像,它的波段为 B2(蓝

波段)、B3(绿波段)、B4(红波段)、B8(近红外波段)。影像数据下载自网站 <https://earthexplorer.usgs.gov>。对梨树县来讲,共需要 4 景影像来覆盖整个研究区。考虑到作物生长季是 5-10 月,剔除云量过多、不同景影像拼接色差大的数据,最终使用的影像信息如表 1 所示。此外,从上述网站上下载的影像是只经过几何精校正的 L1C 级数据,还需进行辐射定标和大气校正。本研究基于欧洲航天局发布的开源软件 SNAP 软件对影像进行辐射定标和大气校正,以生成 L2A 级数据,然后进行波段合成、拼接、裁剪处理,以获取研究区影像。

1.2.2 数字高程数据 本研究还使用了数字高程数据(Digital Elevation Model, DEM),用于计算坡度信息,进行研究区内山区面积的提取。数据下载自地理空间数据云平台(<http://www.gscloud.cn>)。

表 1 本研究使用的 Sentinel-2 影像数据列表

Table 1 List of Sentinel-2 image data used in this study

影像获取时间	成像卫星	云量(%)	影像质量
2018-05-04	Sentinel-2A	0	优
2018-05-24	Sentinel-2A	0	优
2018-08-02	Sentinel-2A	0	优
2018-09-16	Sentinel-2B	5	良
2018-10-01	Sentinel-2A	0	优
2018-11-05	Sentinel-2B	2	良

1.2.3 地面调查数据 地面调查在 2018 年 8 月 5-10 日开展,借助当地农机技术推广站工作人员的经验,选择调查路线以使考察经过最多的植被类型。调查时,基于选定的路线,随机选择样点,记录样点的植被类型、经纬度等信息并拍照。考察中共获取玉米样点 245 个,大豆样点 38 个,水稻样点 64 个,蔬菜样点 68 个,林地样点 56 个,其他样点(建筑、道路、水体等)62 个。所有样点的位置分布如图 1 所示。

1.3 数据分析方法

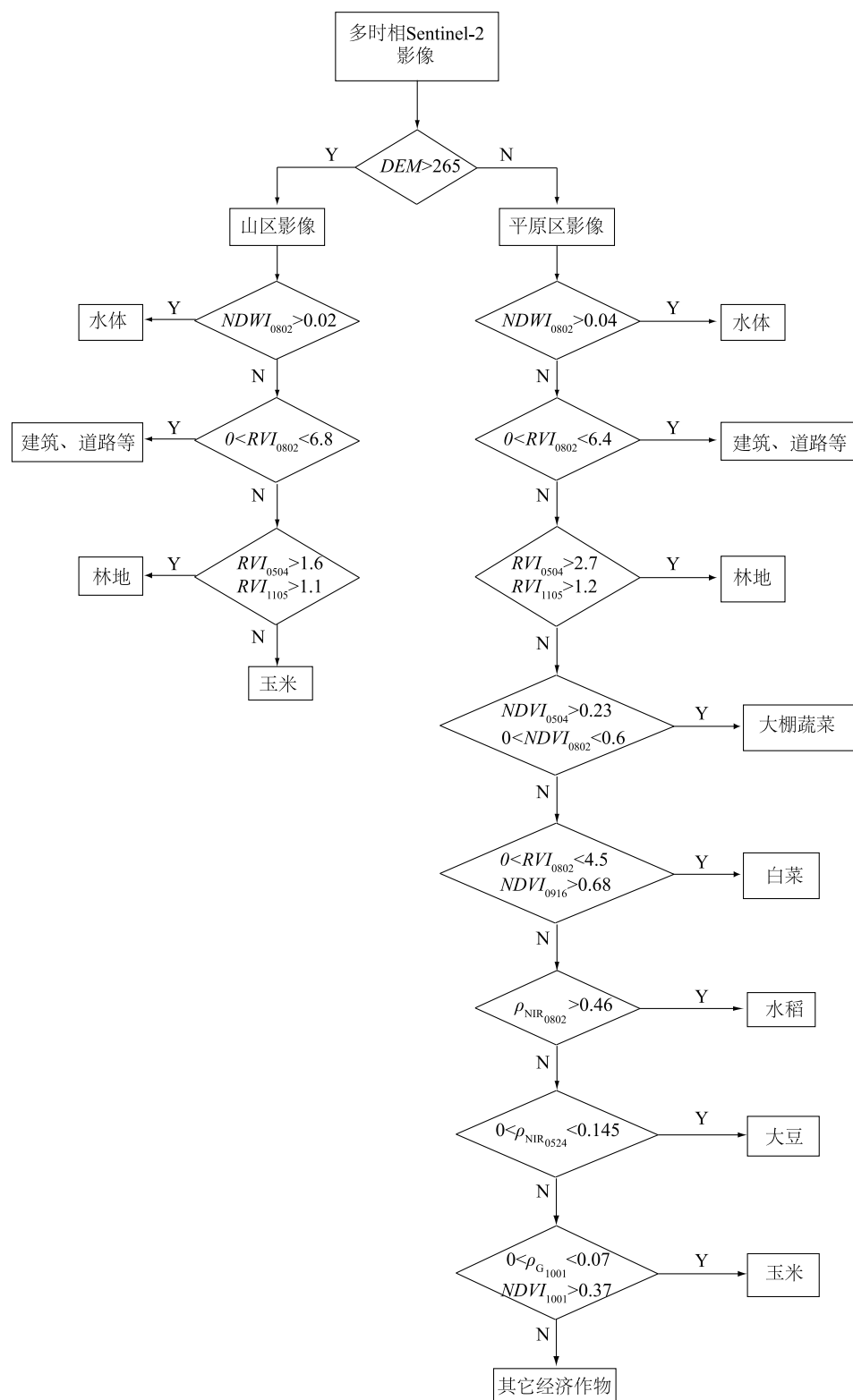
影像分类时,采用了决策树分类、最大似然分类和支持向量机分类 3 种常用的分类方法。首先,将所有采集到的样点随机选择 2/3 作为训练样本,1/3 作为验证样本,然后基于训练样本分别建立 3 种方法下的分类模型,最后基于模型来进行分类并利用验证样本对分类结果进行验证,以比较 3 种方法的优劣。

1.3.1 基于决策树的分类 决策树是一种由内部节点判断属性,分支代表结果输出,叶节点代表类别的树状结构。其分类规则是实现决策树分类的关键,而获取分类规则主要通过专家先验知识、数理归纳统计或者根据 CART、ID3、C4.5 等算法自动获取^[16]。本研究基于训练样本中不同植被类型波段反射率和各种光谱指数的差异,归纳统计确定它们之间区分的阈值,从而基于阈值构建决策树分类模型。

由于梨树县东南部为丘陵区,山区地物光谱特征与平原会有所差异,因此在分类过程中首先基于 DEM 数字高程数据计算坡度,然后利用坡度将研究区分为山区与平原区。根据水体、建筑、道路、林地等非农业用地光谱特征的差异,在山区与平原区中分别提取出这些非农业用地面积。最后,再根据山区与平原区主要作物的物候特征,找出它们之间的光谱特征差异,提取出区域内的各种作物面积,并根据研究目的,通过合并将所有地物分为玉米、大豆、水稻、蔬菜、其他经济作物、非农业用地共 6 类。

技术路线图如图 2 所示,各类地物提取流程详述如下。首先提取非农业用地的面积,非农业用地包括水体、建筑、道路和林地。对于非农业用地,山区与平原区的提取流程是一样的,但所用阈值有所不同。首先,利用归一化水体指数($NDWI$)^[17]将水体面积提取出来,其计算过程如公式(1)所示。其中,在山区选择的阈值为 0.02,在非山区选择的阈值为 0.04。其次,利用比值植被指数(RVI)^[18]提取建筑用地、道路的面积,其计算过程如公式(2)所示。分析不同地物的 RVI 时序曲线,发现建筑用地、道路等的 RVI 值比较低,且不会随时间有太大波动(图 3)。8 月,大部分植被都处于生长旺盛阶段, RVI 值较高,与建筑用地、道路等这些非植被地物相比,差异很大。因此,根据 8 月份的 RVI 信息可以将建筑用地、道路的面积提取出来。其中,在山区选择的阈值为 $0 < RVI_{0802} < 6.8$ (RVI_{0802} 表示 8 月 2 日 RVI),在非山区选择的阈值为 $0 < RVI_{0802} < 6.4$ 。最后,同样基于 RVI 数据提取林地面积。由图 3 可以发现,林地的 RVI 值在 5 月初和 10 月下旬与其他地物类型有较大差异,这主要是因为 5 月初大部分作物刚刚播种,而 10 月下旬都已收获完毕,2 个时期的作物类地块均无植被覆盖,而林地在这 2 个时期还有植被特征。因此,选择这 2 个时期的 RVI 值来提取林地面积。山区林地的分类阈值为 $RVI_{0504} >$

1. 6、 $RVI_{1105} > 1.1$, 非山区林地的分类阈值为 $RVI_{0504} > 2.7$ 、 $RVI_{1105} > 1.2$ 。



DEM: 数字高程; NDWI: 归一化水体指数; RVI: 比值植被指数; NDVI: 归一化植被指数; ρ_{NIR} : 近红外波段反射率; ρ_G : 绿波段反射率。

图2 基于决策树分类的技术路线图

Fig.2 Technical route map based on decision tree classification

$$NDWI = \frac{\rho_G - \rho_{NIR}}{\rho_G + \rho_{NIR}} \quad (1)$$

其中, ρ_G 代表绿波段反射率, ρ_{NIR} 代表近红外波段反射率。

$$RVI = \frac{\rho_{NIR}}{\rho_R} \quad (2)$$

其中, ρ_R 代表红波段反射率, ρ_{NIR} 代表近红外波段反射率。

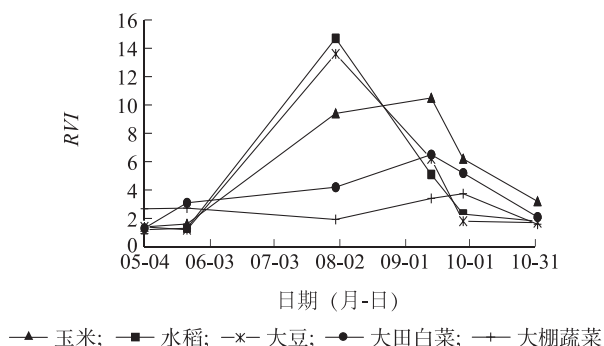


图3 不同地物 RVI 时序曲线

Fig.3 Ratio vegetation index (RVI) time-series curve for different ground objects

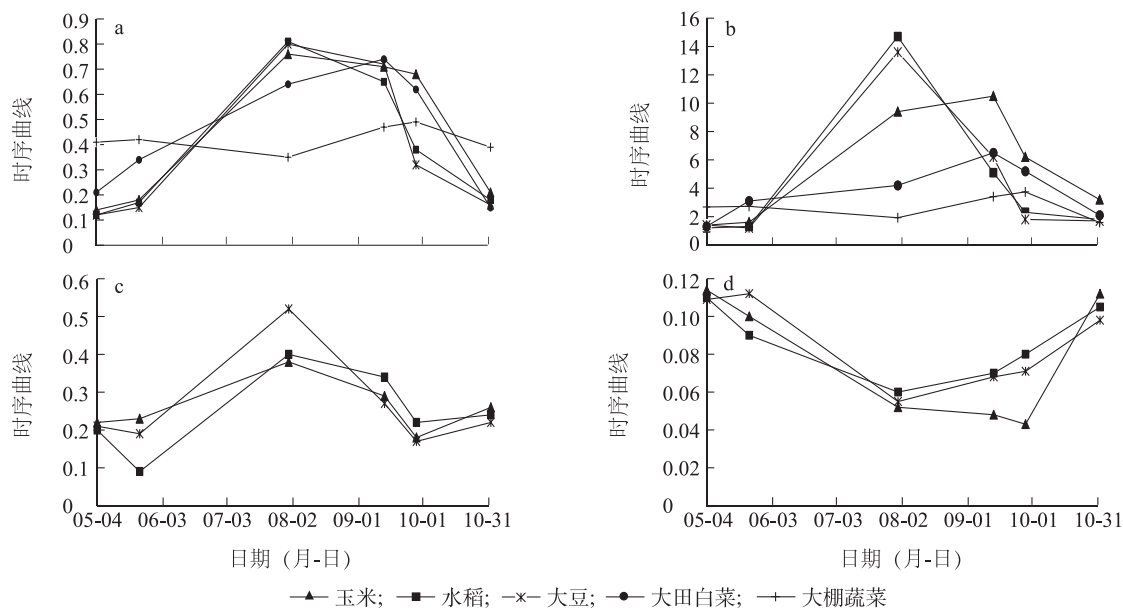
然后提取农业用地的面积,农业用地即种植作物的地块,梨树县主要种植的作物有玉米、水稻、大豆和蔬菜。结合作物的物候变化分析其光谱特征的时序变化,可找出不同作物的分类依据。表2为它们的物候期。由于山区地形崎岖复杂,不便运输,且土壤类型单一,梨树县山区的主要作物只有玉米。因此,在山区,林地面积提取后,剩下的植被区划分为玉米地。在非山区,玉米、水稻、大豆、蔬菜等4种作物都有。首先,提取蔬菜种植面积。蔬菜分为大棚蔬菜和大田蔬菜。大棚蔬菜种植种类多样,而大田蔬菜以白菜为主。本研究基于归一化植被指数(NDVI)^[19]提取大棚蔬菜的种植面积,其计算过程如公式(3)所示。通过对比大棚蔬菜与其他作物 NDVI 特征曲线(图4a),发现大棚蔬菜的 NDVI 值较为平均,一直处于0.35~0.50,并在5月时其 NDVI 在所有作物中处于最高,8月时处于最低。这是由于5月除大棚蔬菜外大部分作物都刚播种,而此时大棚蔬菜相较于这些植被特征不明显的地块对红外光的反射能力较强,对红光的反射能力较弱,则 NDVI 较高。8月是大部分作物植被特征最明显的时候,NDVI 值

比较高,且与大棚蔬菜的 NDVI 差值较大。所以最终确定大棚蔬菜的分类阈值为 $NDVI_{0504} > 0.23$, $0 < NDVI_{0802} < 0.60$ 。另外,通过分析大田作物 NDVI 和 RVI 的特征曲线(图4a、图4b)可发现,大田白菜的 NDVI 和 RVI 在8月初都要低于其他作物,到了9月中旬其 NDVI 和 RVI 会显著升高。这是因为8月初大田白菜刚刚播种,而其他作物已经处于生长较旺盛的阶段,因此其 NDVI 和 RVI 值与其他作物差异较大;9月中旬大田白菜生长至结球期,则 NDVI 和 RVI 会显著升高。对比 NDVI 和 RVI 的时序特征曲线,选择差值较大的特征作为大田白菜的分类依据(图4a、图4b)。因此,选择大田白菜的分类阈值为 $0 < RVI_{0802} < 4.50$ 、 $NDVI_{0916} > 0.68$ 。提取出蔬菜种植面积后,玉米、水稻、大豆的分类主要基于它们物候特征的差异,通过比较它们绿波段反射率 ρ_G 和近红外波段反射率 ρ_{NIR} 的变化(图4c、图4d)将它们的种植面积分别提取出来。通过比较发现5月下旬水稻的 ρ_{NIR} 是所有作物中最低的,这是由于此时水稻刚插秧,水田对红外光的反射能力较弱。因此,选择水稻的分类阈值为 $0 < \rho_{NIR_{0524}} < 0.145$;大豆的 ρ_{NIR} 在8月初处于峰值,且是几种作物之中最高的。这是因为大豆7月下旬进入结荚期,叶绿素的吸收能力逐渐增强,因此对红外光的反射能力增强。因此,选择大豆的分类阈值为 $\rho_{NIR_{0802}} > 0.460$;10月初玉米的 ρ_G 要低于水稻和大豆,且由图4a作物 NDVI 特征曲线可看出此时玉米的 NDVI 值最高。这是因为10月除玉米外大部分作物已经收获完毕,只有玉米还具有植被特征。因此,选择玉米的分类阈值为 $0 < \rho_{G_{1001}} < 0.07$ 、 $NDVI_{1001} > 0.37$ 。

$$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R} \quad (3)$$

其中, ρ_R 代表红波段反射率, ρ_{NIR} 代表近红外波段反射率。

1.3.2 基于最大似然的分类 最大似然分类又称贝叶斯分类,它的原理是根据贝叶斯判决准则进行分类,即假设所有训练样本的分布函数都为正态分布,首先计算每一个待分类的像元归属于每一种训练样本的归属概率,然后将该像元分到归属概率最大的一类样本中^[20]。最大似然法具有简单便捷、理论依据充足的优点。但是它的应用范围较窄,一般只适用于波段较少的数据,且它对训练集大小要求也较高。



a:NDVI 时序曲线;b:RVI 时序曲线;c:近红外波段反射率(ρ_{NIR})时序曲线;d:绿波段反射率(ρ_G)时序曲线。

图 4 不同作物光谱特征曲线

Fig.4 Spectral characteristics curve of different crops

表 2 梨树县主要农作物物候期

Table 2 Phenophase of main crops in Lishu County

农作物	物候变化
玉米	5 月 1 日左右播种,经历拔节期、抽穗开花期、灌浆期、成熟期,10 月 1 号左右开始收获
水稻	5 月中旬左右播种,经历分蘖期、抽穗期、乳熟期、灌浆期、成熟期,大约 10 月上旬收割
大豆	4 月下旬左右播种,经历分枝期、开花期、结荚期、鼓粒期、成熟期,9 月中下旬左右收获
蔬菜	大田主要以白菜为主,8 月初播种,经历莲座期、结球期、成熟期,10 月中下旬收获;冬季以大棚蔬菜为主,且种植种类多

在前面决策树方法中,已详细描述作物物候特征与光谱信息间的关系,以及由此带来的最佳区分各种作物的时间与所用光谱特征参数,这些规律同样可在其他分类方法中应用。因此,为了保证 3 种方法数据源相同,在基于最大似然分类中,本研究首先将决策树分类法中涉及的坡度和光谱特征等 ($Slop$ 、 $NDWI_{0802}$ 、 RVI_{0504} 、 RVI_{0802} 、 RVI_{1105} 、 $NDVI_{0504}$ 、 $NDVI_{0916}$ 、 $NDVI_{1001}$ 、 $\rho_{NIR0524}$ 、 $\rho_{NIR0802}$ 、 ρ_{G1001}) 共 11 个数据作为波段合并到同一影像中,然后利用训练样本集数据基于最大似然法训练得到分类模型,最后基于分类模型进行分类获得相关分类结果。

1.3.3 基于支持向量机的分类 支持向量是指距离分类超平面近的那些点,支持向量机的基本思想就是在样本空间中构建出最优分类超平面,使得支

持向量距离该分类超平面的间隔最大,这个最大间隔代表了该分类超平面对两类数据的区分度强,不易错分。最优分类超平面的方程和约束条件如公式(4)、公式(5)所示。支持向量机具有计算复杂度低、稳定性较强、抗噪性强等优点。其模型思路理论上会获得很高的分类精度,但实际使用中获取高精度需要取决于很多因素,如数据的预处理精度、训练集的大小、特征参数及核函数的选择等^[21]。

同最大似然法的分类过程相同,本研究首先合并坡度数据和相应光谱特征为同一影像,然后再基于训练样点通过支持向量机训练得到分类模型,最优基于分类模型进行分类获得相关分类结果。

$$\omega^T x_i + b = 0 \tag{4}$$

$$y_i(\omega^T x_i + b) \geq 1 - \xi_i \tag{5}$$

其中, x_i 为输入模式, y_i 为目标输出, ω 为权值向量, b 为偏置, ξ_i 为线性不可分条件下的松弛变量。

1.3.4 精度验证 本研究使用地面调查数据中各类作物的验证样本对分类结果进行验证,计算混淆矩阵、每一种作物的制图精度和用户精度以及总体分类精度和 Kappa 系数。精度验证的计算公式如表 3 所示,其中 k 代表混淆矩阵行列数即分类的类别数量, X_{ii} 代表混淆矩阵中对角线上的值即正确分类的像元数, N 代表验证的像元总数, X_i 代表混淆矩阵第 i 行元素相加之和, X_j 代表混淆矩阵第 j 列元素相加之和。

表 3 精度验证的计算公式

Table 3 Formulas for accuracy verification

类型	计算公式
制图精度 (PA)	$PA = \frac{X_{ii}}{X_j} \times 100\%$
用户精度 (UA)	$UA = \frac{X_{ii}}{X_i} \times 100\%$
总体精度 (OA)	$OA = \sum_{i=1}^k \frac{X_{ii}}{N} \times 100\%$
Kappa 系数 (K)	$K = \frac{N \sum_{i=1}^k X_{ii} - \sum_{i=1}^k X_i X_j}{N^2 - \sum_{i=1}^k X_i X_j}$

2 结果与分析

2.1 分类结果精度评价

基于验证样本集,决策树法、最大似然法和支持向量机法的验证结果如表 4~表 6 所示。从误差矩阵看,在所有作物中,玉米、水稻在 3 种分类方法中的分类精度最高。其中,玉米的制图精度分别为 96.34% (决策树法)、93.90% (最大似然法)、96.43% (支持向量机法),用户精度分别为 96.34% (决策树法)、93.90% (最大似然法)、92.05% (支持向量机法)。这是由于其种植面积大,获取的样本点最多,较多的样本点保证了玉米具有很好的制图和用户精度。水稻的制图精度分别为 100.00% (决策树法)、95.24% (最大似然法)、95.24% (支持向量机法),用户精度分别为 95.45% (决策树法)、95.24% (最大似然法)、95.24% (支持向量机法)。水稻的分类精度高,一方面是因为 5 月下旬时水田特征很明显,其近红外波段的反射率与其他地物差别大,其种植面积的提取效果较好;另一方面,梨树县的水稻多为集中种植,各田块光谱特征差异较小。蔬菜种类繁多,包括大棚蔬菜和大田蔬菜,分类精度中等。其中,制图精度分别为 83.33% (决策树法)、84.21% (最大似然法)、82.61% (支持向量机法),用户精度分别为 86.96% (决策树法)、88.89% (最大似然法)、100.00% (支持向量机法)。大豆的分类精度最低。其中,制图精度分别为 83.33% (决策树法)、70.00% (最大似然法)、66.67% (支持向量机法),用户精度分别为 83.33% (决策树法)、63.64% (最大似然法)、66.67% (支持向量机法)。这是因为大豆的种植区域比较分散,不易寻找,在采集地面样点时获取的样点比较少,缺乏足够的样本

来训练导致的。

从 3 种分类方法的总体分类精度看,决策树法的总体分类精度为 93.53%, Kappa 系数达到 0.890 6;最大似然法的总体分类精度为 90.91%, Kappa 系数为 0.838 5;支持向量机法的总体分类精度为 91.43%, Kappa 系数为 0.849 8。3 种方法中决策树法的分类精度最高。

表 4 决策树法分类结果的误差矩阵

Table 4 Error matrix of classification results by decision tree method

作物类别	玉米	水稻	大豆	蔬菜	总计	用户精度 (%)	制图精度 (%)
玉米	79	0	1	2	82	96.34	96.34
水稻	0	21	0	1	22	95.45	100.00
大豆	1	0	10	1	12	83.33	83.33
蔬菜	2	0	1	20	23	86.96	83.33
总计	82	21	12	24	139		

表 5 最大似然法分类结果的误差矩阵

Table 5 Error matrix of classification results by maximum likelihood method

作物类别	玉米	水稻	大豆	蔬菜	总计	用户精度 (%)	制图精度 (%)
玉米	77	1	3	1	82	93.90	93.90
水稻	0	20	0	1	21	95.24	95.54
大豆	3	0	7	1	11	63.64	70.00
蔬菜	2	0	0	16	18	88.89	84.21
总计	82	21	10	19	132		

表 6 支持向量机法分类结果的误差矩阵

Table 6 Error matrix of classification results by support vector machine method

作物类别	玉米	水稻	大豆	蔬菜	总计	用户精度 (%)	制图精度 (%)
玉米	81	1	4	2	88	92.05	96.43
水稻	0	20	0	1	21	95.24	95.24
大豆	3	0	8	1	12	66.67	66.67
蔬菜	0	0	0	19	19	100.00	82.61
总计	84	21	12	23	140		

2.2 作物种植结构分析

基于决策树的分类结果如图 5 所示。对梨树县种植结构进行分析,发现粮食作物的种植面积远远超过经济作物(表 7),这是因为梨树县的自然条件适于旱作农业,尤其适宜种植玉米。因此粮食作物

的种植面积较大,是中国的重要粮食基地。从空间分布看,玉米总种植面积为235 619 hm²,占作物总种植面积的 84.30%。由于中部地区是平原,土地平坦肥沃,因此中部万发镇、泉眼岭乡、金山乡、白山乡等地区大面积种植玉米。水稻种植面积占作物总种植面积的 3.76%,约为10 515 hm²,其主要分布在北部地区的小宽镇、孤家子镇,这是因为梨树县北部地区是风沙盐碱地,地势低洼,适宜水稻生长。由于白菜是喜湿作物,因此白菜也主要分布在水源较多的小宽镇、孤家子镇及小城子镇。大棚蔬菜集中在中心城区梨树镇周围,主要为了交通运输方便。蔬菜的总种植面积占梨树县作物总种植面积的 3.40%,约为9 507 hm²。南部十家堡镇和孟家岭镇地块少,多为山地丘陵地带,因此山林灌木众多,但也种植少量玉米。西部地区喇嘛甸镇、林海镇、刘家馆子镇、四棵树乡经济作物的种类繁多,这是因为西部地区土壤类型丰富,适宜种植大豆等经济作物。

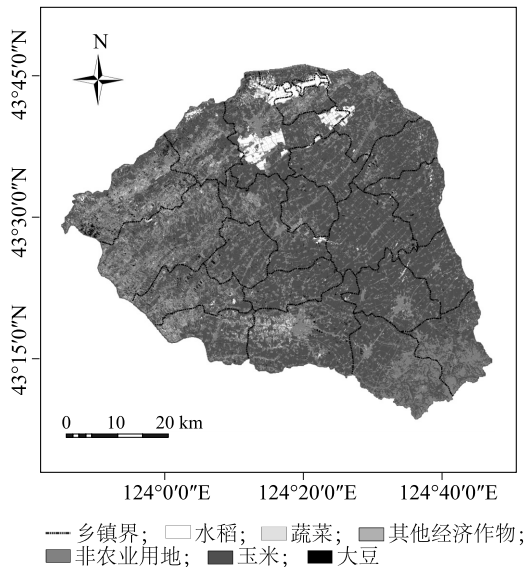


图5 2018年梨树县农作物决策树法分类结果
Fig.5 Crop classification results by decision tree method for Lishu County in 2018

表7 2018年梨树县农作物种植结构
Table 7 Crop planting structure for Lishu County in 2018

农作物	面积(hm ²)	所占比重(%)
玉米	235 619	84.30
水稻	10 515	3.76
大豆	5 521	1.97
蔬菜	9 507	3.40
其他经济作物	18 338	6.57

3 讨论

基于多时相 GF-WFV 和 GF-2PAN 数据,Zhang 等^[22]采用 Sobel 边缘检测和 Hough 变换纹理分析识别玉米,精度高达 90%;赵建鹏等^[23]使用多时相的 ZY-3 和 GF-2 数据,通过构建 NDVI 差值模型提取固安县柳泉镇冬小麦种植面积,精度达到 95.54%;杜保佳等^[24]基于 Sentinel-2A 影像对北安市作物进行分类,引入 NDVI 时间序列特征后,使用多时相数据的分类精度相较于使用典型单时相数据提高 7.7%。不同作物不同时期长势存在差异,可以通过它们在多时相影像光谱特征的差异来对其进行识别,从而提高作物的分类精度,而使用单景卫星影像时,往往会出现云量多、植被长势无差异等现象,使得分类精度降低。本研究同样使用了多时相遥感数据,通过分析不同地物光谱特征的时序变化,分别使用决策树法、最大似然法、支持向量机法 3 种分类方法,最终得到 93.53%、90.91%、91.43% 的分类精度,与以上基于其他高分影像的分类精度差不多。

本研究在提取非农业用地和大田白菜的种植面积时,比较了 RVI 和 NDVI 2 种光谱指数的分类效果,发现在高植被覆盖度下,RVI 的提取效果要优于 NDVI。比如在利用夏季影像提取非农业用地面积时,RVI 的效果要优于 NDVI,能更好地体现出地块的轮廓。在利用 9 月中旬影像提取大田白菜等排列不紧密、覆盖度较低的植被面积时,NDVI 的提取效果要优于 RVI,而此时使用 RVI 提取大田白菜种植面积时会出现大量的漏分。这印证了钱铭杰^[25]在比较 RVI 与 NDVI 在植被面积提取中的研究结果,当植被覆盖度高时,RVI 对植被的敏感性要高于 NDVI,反之则 NDVI 的敏感性更高。因此,为提高分类精度,在提取不同地物面积时应结合实际情况,选择不同的植被指数。

本研究中决策树法的分类精度要高于最大似然法、支持向量机法这 2 种监督分类算法,这是由于决策树法相对于其他算法不需要假设先验概率分布,具有更好的灵活性和鲁棒性,且对于数据集的离散性和规则性的要求都较低。王姝力等^[26]基于光学数据使用 CART 自动决策树法识别海冰类型的精度要高于使用最大似然法的识别精度。许童羽等^[27]使用 CART 决策树和 BP 神经网络相结合的方法提取粳稻种植面积,相较于单一通过 BP 神经网络的

方法,精度提高了 1.70%。本研究与这些研究结果类似,说明决策树法的实用性和准确性。

本研究设定的决策树法作物分类阈值目前只基于 2018 年数据进行了验证,其是否适合在其他年份或其他地区应用,还有待进一步验证。

4 结论

本研究基于多时相 Sentinel-2 影像和地面采集样点数据,分别通过决策树法、最大似然法、支持向量机法 3 种分类方法实现了梨树县作物种植面积自动提取,并对它们的分类效果进行了对比评价,并在此基础上分析了梨树县作物的种植结构空间分布情况。结果表明,决策树法总体分类精度最高为 93.53%,Kappa 系数达到 0.890 6。从种植结构上看,梨树县以种植粮食作物为主,经济作物为辅。其中,玉米在全县各区域都有,但集中在中部地区;水稻多种植在北部地区;大豆等经济作物主要种植在西部地区。耦合时序 Sentinel-2 影像和本研究所构建的决策树自动分类方法能很好地对梨树县作物进行分类,从而达到客观评价其种植结构的目的。

参考文献:

- [1] 胡琼,吴文斌,宋茜,等.农作物种植结构遥感提取研究进展[J].中国农业科学,2015,48(10):1900-1914.
- [2] 唐华俊,吴文斌,杨鹏,等.农作物空间格局遥感监测研究进展[J].中国农业科学,2010,43(14):2879-2888.
- [3] 张养贞,张树文,常丽萍,等.县级玉米遥感估产实验及其效果研究[J].地理科学,1995(2):144-153.
- [4] 张荣群,王盛安,高万林,等.基于时序植被指数的县域作物遥感分类方法研究[J].农业机械学报,2015,46(S1):246-252.
- [5] 刘珺,田庆久,黄彦,等.利用多时相 HJ 卫星 CCD 遥感影像提取嘉祥县秋收作物[J].遥感信息,2012,27(2):67-70.
- [6] NASRALLAH A, BAGHDADI N, MHAWEJ M, et al. A novel approach for mapping wheat areas using high resolution sentinel-2 images[J].Sensors, 2018,18:7.
- [7] SONOBE R, YAMAYA Y, TANI H, et al. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning[J].Journal of Applied Remote Sensing,2018,12(2):26-45.
- [8] ZHU J, PAN Z W, WANG H, et al. An improved multi-temporal and multi-feature tea plantation identification method using Sentinel-2 imagery[J].Sensors,2019,19:9.
- [9] CORENTIN B, ADRIEN M, PETER G, et al. Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery[J].Biotechnologie, Agronomie, Société et Environnement,2018,22:3.
- [10] 王大钊,王思梦,黄昌. Sentinel-2 和 Landsat8 影像的四种常用水体指数地表水体提取对比[J].国土资源遥感,2019,31(3):157-165.
- [11] 刘怀鹏,安慧君.基于 Sentinel-2A 的农田包围型村落提取[J].内蒙古农业大学学报(自然科学版),2019,40(3):41-45.
- [12] 王祁春,张柏,张树文,等.玉米长势区域分异遥感监测——以松嫩平原玉米遥感估产实验区(梨树县)为例[J].遥感信息,1994(4):20-23.
- [13] 平跃鹏,臧淑英.基于 MODIS 时间序列及物候特征的农作物分类[J].自然资源学报,2016,31(3):503-513.
- [14] 张晓萌,刘建祥,温馨,等.基于遥感的植被覆盖度和水土流失信息提取[J].水土保持,2017,5(4):21-28.
- [15] 范唯唯. Sentinel-2B 卫星发射成功[J].空间科学学,2017,37(4):371-372.
- [16] 张润雷.基于决策树的遥感图像分类综述[J].电子制作,2018(24):16-18,55.
- [17] AHMED K R, AKTER S. Analysis of landcover change in south-west bengal delta due to floods by NDVI, NDWI and K-Means cluster with landsat Multi-spectral surface reflectance satellite data[J]. Remote Sensing Applications: Society and Environment, 2017(8):168-181.
- [18] 刘人午.基于 LANDSAT 卫星数据对内江主城区城市绿化监测研究[J].内江科技,2017,38(12):38-41.
- [19] TIAN Y C, BAI X Y, WANG S J, et al. Spatial-temporal changes of vegetation cover in guizhou province, southern China[J].Chinese Geographical Science,2017,27(1):25-38.
- [20] 王婷婷,李山山,李安,等.基于 Landsat 8 卫星影像的北京地区土地覆盖分类[J].中国图象图形学报,2018,20(9):1275-1284.
- [21] 李平,吴曼乔,曾联明.支持向量机技术在土地利用监测的应用研究[J].测绘通报,2010(8):28-30.
- [22] ZHANG C, JIN H, LIU Z, et al. Seed maize identification based on texture analysis of GF remote sensing data[J].Transactions of the Chinese Society of Agricultural Engineering, 2016,32(21):183-188.
- [23] 赵建鹏,李国洪,金永涛,等.县域尺度上基于多时相影像的冬小麦面积监测[J].河北师范大学学报(自然科学版),2018,42(6):524-530.
- [24] 杜保佳,张晶,王宗明,等.应用 Sentinel-2A NDVI 时间序列和面向对象决策树方法的农作物分类[J].地球信息科学学报,2019,21(5):740-751.
- [25] 钱铭杰. RVI 与 NDVI 在植被信息提取中的应用比较[C]//中国地理信息系统协会中国遥感协会.第七届 ArcGIS 暨 ERDAS 中国用户大会论文集.北京:地震出版社,2006:662-666.
- [26] 王殊力,王志勇,王磊.基于 Landsat-8 和 Sentinel-1A 辽东湾海冰分类研究[J].北京测绘,2019,33(12):1486-1492.
- [27] 许童羽,胡开越,周云成,等.基于 CART 决策树和 BP 神经网络的 landsat 8 影像梗稻提取方法[J].沈阳农业大学学报,2020,51(2):169-176.

(责任编辑:陈海霞)