

构建地块二维表征及 CNN 模型的作物遥感分类

张乾坤^{1,2}, 蒙继华², 任超¹

1. 桂林理工大学 测绘地理信息学院, 桂林 541004;

2. 中国科学院空天信息创新研究院 数字地球重点实验室, 北京 100101

摘要: 本文旨在研究基于地块数据约束的深度学习模型的分类特征表示方法, 以识别不同作物在不同时相上光谱差异从而对作物类型进行分类。通过 Google Earth Engine 平台获取作物生育期内全部 Landsat 8 影像, 利用其质量评定波段完成研究区无云时相及区域上的地块统计, 提取地块级别的各波段反射率均值按照时相顺序及波长进行排列, 构建波谱、时相二维特征图作为该地块的抽象表示。通过构建相对最优的卷积神经网络 CNN (Convolutional Neural Network) 结构完成对特征图的分类, 从而完成对地块的分类。构建 CNN 模型并不需要手工特征和预定义功能的需求, 可完成提取特征并遵循端到端原则进行分类。将该模型的分类结果与其他最为常用机器学习分类器进行了比较, 获得了优于常用遥感分类算法的分类精度。结果表明地块数据的加入可以有有效的缩减计算规模并提供了准确的分类边界。所提出方法在地块特征表示及作物分类中具有突出的应用潜力, 应视为基于地块的多时相影像分类任务的优选方法。

关键词: 遥感, 作物分类, CNN, 特征图, 二维表征, 时序影像, 地块数据

引用格式: 张乾坤, 蒙继华, 任超. 2022. 构建地块二维表征及 CNN 模型的作物遥感分类. 遥感学报, 26(7): 1437-1449

Zhang Q K, Meng J H and Ren C. 2022. Crop classification based on two-dimensional representation and CNN model from remote sensing. National Remote Sensing Bulletin, 26(7): 1437-1449 [DOI: 10.11834/jrs.20219432]

1 引言

通过农业遥感监测和调查能够及时准确地获取作物种植面积及时空分布信息, 对农业经营管理、种植业结构调整、国家粮食安全均具有重要意义 (平跃鹏, 2016)。农作物的时空分布是农业生产过程中的基本信息, 也是掌握植物生长状况, 指导农业生产, 开展其他农业遥感应用的重要依据 (陈仲新等, 2016)。如何通过遥感手段对农作物的类型精准识别仍然是农业遥感研究中最活跃的议题之一 (Onojeghuo等, 2018)。遥感识别的本质是从传感器对地观测的影像到代表用户感兴趣的地面覆盖类型的标签字段的映射 (Li等, 2019)。监督分类是农作物遥感识别的主要方法, 依赖于有足够的类标签已知的可用像元来训练分类器。通过训练完成对分类器所需参数的估计, 以便能够识别和标记其他未标记像元 (Seal等,

2020)。根据分类对象是否为像元而又可以分为逐像元分类和面向对象分类 (Rahman 和 Saha, 2008; Blaschke, 2010)。面向对象分类采用了由邻近像元所构成的图斑作为分类单元, 图斑对象不仅包含原始的像元属性, 还可以加入全局性的光谱特征统计变量以及局部性的空间特征 (Zhang等, 2017), 可以克服传统逐像元进行作物分类精度较低、分类结果常常是椒盐图像等问题。作物遥感分类与其他的类型的地物遥感分类的主要区别在于不同农作物都表现出各自典型的物候特征 (Cai等, 2018), 其光谱在不同季节有着明显的变化, 并且在区域内同一种作物的生长发育规律相对稳定 (李中元等, 2019)。越来越多的研究显示增加植物生长变化更加敏感的波段和利用多时相的遥感影像可以提高作物的识别能力 (Sun等, 2019)。遥感技术的发展提供了便利, 对地观测卫星搭载的传感器空间、波谱、时间分辨率逐渐提升 (史

收稿日期: 2019-12-18; 预印本: 2020-05-19

基金项目: 高分辨率对地观测系统重大专项 (编号: 09-Y20A05-9001-17/18, 30-Y20A03-90030-17/18)

第一作者简介: 张乾坤, 研究方向为农作物遥感分类。E-mail: zhang_q_k@163.com

通信作者简介: 蒙继华, 研究方向为作物遥感监测及精准农业遥感应用。E-mail: mengjh@radi.ac.cn

舟等, 2015)。采用多时相影像和加入敏感波段带来精度提升的本质是参与分类的特征得到扩充。这种扩充不仅仅是原始的波谱特征的增加, 也表明研究者有更多的选择来构造衍生特征 (Su 等, 2018)。分类特征的增加使得模型将更加复杂, 其泛化能力下降, 而且特征变量过多会引起“维数灾难”, 识别精度反而会降低 (Cavallaro 等, 2015)。尽管这也使得特征提取、特征选择等手段有着更多的选择方案 (Zhang 等, 2012, 2017; Valero 等, 2016), 但在实践中, 要找到有效且合适的方案并非易事。手工设计的特征工程依赖于人的经验和领域知识, 且冗长和耗时的。通常需要人为干预来应对给定条件的变化。人类知识很难同时考虑到类间相似性、类内变异性、大气条件和光散射机制等复杂因素 (Zhong 等, 2019)。预先定义的模型和相关的数学假设的固定形式限制了处理不同模式的灵活性。

神经网络具有近似任意非线性函数的能力, 通过简单组合但非线性的模块进行多层表征学习 (LeCun 等, 2015)。经典的神经网络称为多层感知器 MLPs (MultiLayer Perceptrons), 可以应用于各类的数据预测问题。MLPs 近似一个从输入变量到输出变量的映射函数 (Hinton 等, 2012)。这种通用能力对于监督分类任务是有价值的, 一方面是其神经网络对输入数据和映射函数中的噪声具有很强的鲁棒性, 甚至可以在缺失值存在的情况下支持学习和预测。另一方面, 神经网络对映射函数没有很强的假设, 容易学习线性和非线性关系, 支持映射函数中任意定义但固定数量的输入和输出。这意味着多变量输入, 可以指定任意数量的输入特征, 为多元预测提供直接支持。卷积神经网络 CNN (Convolutional Neural Network) 通过卷积核作为中介, 在每一图像层共享卷积核内的参数, 图像通过卷积操作保留原始位置关系形成下一图像层 (LeCun 等, 2015)。共享权值 (卷积核) 带来的直接好处是减少网络各层之间的连接, 极大的限制参数的数量, 克服了训练 MLPs 梯度消失和梯度爆炸的技术问题 (Shin 等, 2016)。CNN 在遥感领域已开展多方面的应用研究 (Wang 等, 2017; Shao 等, 2018)。利用卷积机制结合堆叠自编码高效的特征提取及特征表示能力一方面降低数据复杂性, 另一方面提升数据表征的有效性 (Zabalza 等, 2016; Han 等, 2016; Lu 等, 2018)。

该方法多用于高光谱影像的降维以及雷达影像的降噪。通过构造更为复杂的卷积层 (Wei 等, 2018)、构建多流网络结构 (Anwer 等, 2018)、进行多模型的决策融合 (Zhang 等, 2018) 等方式增加非线性映射的复杂程度来提升影像的分类精度。这种通过延扩网络结构的方式虽然能应对更为复杂的分类情景, 但同时其要求更多的样本来训练其内部参数而使得应用受限。 (Long 等, 2015; Audebert 等, 2017) 采用全卷积神经网络 (FCN) 络始终保持影像的二维结构产生与原影像同尺寸的类别输出。但其类别边界分类模糊, 与实际有较大差距。虽然 (Bittner 等, 2017; Guo 等, 2018) 利用条件随机场 CRF (Conditional Random Field) 做后处理细化区域边界, 但问题并没有得到完全解决。另外, CNN 通常用于空间和光谱领域, 但在遥感研究中几乎不涉及时间维度。

调用已有地理信息作为条件约束参与遥感应用之中也逐步成为遥感应用的发展趋势。地块数据在遥感分类中起着一种地理约束功能。其作为同一权属且完整封闭的农田, 提供边界、位置、面积等地理信息, 具有长期的稳定性 (韩衍欣, 2018)。很自然的与面向对象分类进行联系, 地块可以作为特殊对象参与分类, 相较于利用影像分割而获得的图斑对象, 地块有着更多的优势 (黄启厅等, 2016; 韩衍欣和蒙继华, 2019)。除了其提供准确的分类边界, 还可以禁止了农田外的其他地物类别参与分类。地块数据的加入对于神经网络还提供潜在的优势, 可以有效的缩减神经网络模型计算规模。使得多类深度学习算法如全卷积神经网络 FCN (Fully Convolutional Networks) 等算法并不需要以完成像素到像素的映射。一方面是研究单元数量的大幅减少, 另一方面可以避免重叠卷积的使用而减少过多的冗余计算。训练神经网络的结构复杂性也因此得到极大降低。

本文提取地块全部影像的各波段地表反射率均值, 地块单元的地表反射率在波谱和时相上仍具有二维特征。从而构建二维表征形式的特征图作为该地块的抽象表示。通过构建合适 CNN 模型完成对光谱、时间维度上的复合表征, 完成对所构建特征图的分类, 从而达到对地块的分类。最后对该模型的分类结果与其他一些机器学习分类器进行了比较。期望通过适当的可视化和映射设置有效地评估分类结果。

2 研究区与数据

2.1 研究区介绍

选择新疆生产建设兵团第八师（简称为八师）为研究区。八师位于中国西北地区，地处天山北麓中段，准噶尔盆地南缘，分布在石河子、克拉玛依、沙湾、玛纳斯境内。研究区位置如图1所示。研究区属于典型的温带大陆性气候，气温年差日差较大，平均气温7.5℃—8.2℃，无霜期147—191 d；气候干旱，降水稀少，年降水量180—270 mm，年蒸发在1000—1500 mm；日照充沛，年日照时长在2721—2818 h；垦区内地势平坦，平均海拔约为450 m；田块规整，作物规模种植，

农业机械化程度较高。研究区生育期内可获得无云或少云影像较多，种植结构及种植模式稳定，区域内自然及生产条件差异较小。

受自然和生产条件限制，实验区种植模式稳定，多为一年一熟单作。其主要作物包括棉花、玉米、小麦、辣椒、桃子、葡萄、瓜类等。棉花种植模式属于典型的西北内陆棉区种植模式于4月播种，9月到10月收获。玉米种植区划属于西北灌溉玉米区，多为一年一熟春玉米单作，于4月种植，也存在少量的青贮玉米于6月种植。小麦种植区划属于新疆冬春播秋春性麦区，八师小麦于10月前后播种，6月中下旬收获。此外，该地区还种植较多瓜果蔬菜，葡萄、桃子、辣椒占有较大比重。

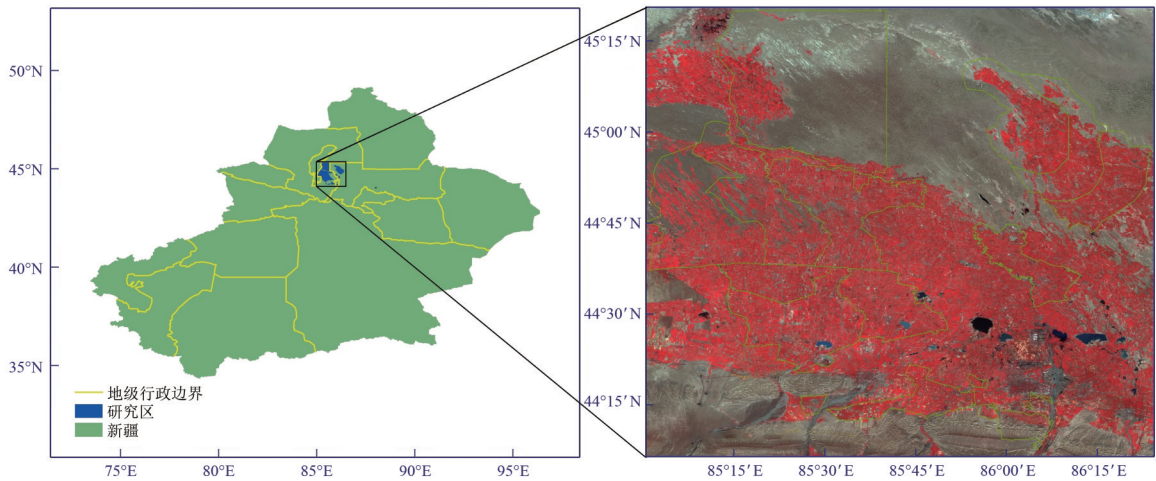


图1 研究区位置及其区域
Fig. 1 Location and area of study area

2.2 遥感影像

采用Landsat 8数据作为本次实验的遥感数据来源。Landsat 8卫星由美国航空航天局（NASA）发射，搭载两个传感器，分别是OLI陆地成像仪（Operational Land Imager）和TIRS热红外传感器（Thermal Infrared Sensor）。Landsat 8在空间分辨率和光谱特性等方面如表1所示，卫星一共有11个波段，波段1—7，9—11的空间分辨率为30 m，波段8为15 m分辨率的全色波段，卫星每16 d可以实现一次全球覆盖。

数据获取于Google Earth Engine平台的USGS Landsat 8表面反射率数据集，该数据集由Google使用USGS提供的Docker镜像利用Landsat 8 OLI/TIRS传感器数据，经过大气校正获得的包含4个

可见光、近红外（VNIR）波段以及2个短波红外（SWIR）波段，经过正射校正后的地表反射率，以及2个热红外（TIR）波段经过正射校正后的辐射亮度。即相对原始数据去除了全色波段、和用于评估数据质量的卷云波段，同时将两个近红外重采样到30 m分辨率。获得覆盖试验区生育期内4—10月全部Landsat 8影像如表2所示。

2.3 地块及样本数据

地块数据是通过人工勾绘完成地块矢量化，采用的是Google Earth 17级数据，空间分辨率为1.19 m，Google Earth影像包含多类遥感平台数据以及一些航拍数据，为多景多源数据拼接且时间并不统一、非现势影像，但其具有较高分辨率，

包含较为准确的地理信息。根据 Google Earth 影像以及高分辨率影像 GF、Sentinel-2 等数据参考勾绘地块矢量，该套地块数据在最初于 2017 年完成。并在此基础上于 2019 年 8 月中旬前往研究区通过进行地面调查以及参考当年高分辨率影像对该套地块数据进行修正。在实验区对地块类别进行地面调查并通过人工解译标注了共计了 2184 个作物样本地块，其中实地调查了 1160 个地块类别。全部地块样本包含了 233 个玉米地块、120 个小麦地块、873 个棉花地块、109 个辣椒地块、128 个包含西瓜、打瓜、甜瓜、哈密瓜在内的瓜类地块、40 个番茄地块、98 个桃树地块、297 个葡萄地块以及 286 个包含荒地、苗圃、树林等林草地块。地块数据及样本数据如图 2 所示。

表 1 Landsat 8 波段配置

Table 1 Landsat 8 band setting

波段序号	波段名称	波长范围/nm	空间分辨率/m
Band1	Coastal(海岸波段)	0.43—0.45	30
Band2	Blue(蓝波段)	0.45—0.51	30
Band3	Green(绿波段)	0.53—0.59	30
Band4	Red(红波段)	0.64—0.67	30
Band5	NIR(近红外波段)	0.85—0.88	30
Band6	SWIR 1(短波红外 1)	1.57—1.65	30
Band7	SWIR 2(短波红外 2)	2.11—2.29	30
Band8	Pan(全色波段)	0.52—0.90	15
Band9	Cirrus(卷云波段)	1.36—1.38	15
Band10	TIRS 1(热红外 1)	10.60—11.19	100
Band11	TIRS 2(热红外 2)	11.50—12.51	100

表 2 研究区作物生育期内影像

Table 2 RS image of crop growth period in the study area

影像日期	云覆盖/%
2019-04-14	2.69
2019-04-30	22.66
2019-06-01	6.58
2019-06-17	3.39
2019-07-03	0.11
2019-07-19	27.55
2019-08-04	81.03
2019-08-20	66.04
2019-09-05	100
2019-09-21	0.46
2019-10-07	0.43

2.4 云掩膜数据

Landsat 8 影像的云掩膜数据来源于 Landsat 8 的质量评定频段。质量评定波段定义某些质量条件的位组合在 PIXELQA 波段中显示为整数值。将像元值所表示的位分解为可理解的条件描述。对 PIXELQA 波段中可能的像元值进行二进制变换后的可以得到影像像元的质量。质量评定波段由 16 位二进制数表示，其二进制位所具有的具体含义说明如表 3 所示，其中位 11—15 未利用。通过 GEE 平台对 PIXELQA 进行波段计算可以获得云掩膜数据。

3 研究方法

本文技术路线如图 3 所示。利用 Landsat 8 质量评定波段获取生育期内全部影像的云掩膜文件。在做地块的统计过程中仅统计无云区域，若地块所在区域均被云覆盖，采用其它影像进行二维插值获得。构建地块特征图作为 CNN 的输入，通过逐步优选模型结构及参数构建相对有效模型。最后完成对测试集分类精度评价及分类结果分析。

3.1 地块的二维表征

地块的表征可以作为地块种植信息的反映，也可以看作地块种植信息被加工的客体。从卫星遥感的角度，地块种植信息的体现是在于对地卫星的观测而获得的遥感影像，更近一步说是影像上各像元反射率值。而区分目标物的属性即特征具有相对性，它总是相对于两个或多个相互比较的目标物而言。因此像元反射率值可以被作为区分种植类别的原始特征存在。表征是目标物信息的呈现方式，保留了主要信息，其具有抽象性，并受到规则的控制。我们把区分地块种植类别的属性按照一定规则进行编码而获得的输出称之为地块表征。在以地块为研究单元，忽略了地块的位置、形状、尺寸等几何信息，其在光谱和时相上仍具有二维特性。可以通过一定的编码方式对二维特性得以保留，通过该方式而获得的输出被称之为地块的二维表征。

地块表征的构建出于作物类型信息在生育期整体性的考虑。不同的地块的表征我们采用相同尺寸的特征图来具体体现。特征图既是表征表示形式也是数据储存形式。特征图是通过地块矢量对影像集内各时相及波段进行区域统计。由于云

覆盖的原因，对于地块统计出来的数字与实际上的存在很大的差距。云覆盖区域的可见光波段远高于实际值。热红外波段低于实际值。对于单一时相而言，云对地块有3种情况，（1）未遮盖，（2）部分遮盖，（3）全部遮盖。对不同的情况分别做不同的处理。对于未遮盖的地块，只需要对地块区域的像素进行统计即可。地块的区域统计最为常用的就是均值统计，一般认为地块像素均

值可以作为地块的整体反射率。一个地块内的作物类别相同，种植条件接近，在地块内的差异小，故而对于部分遮盖的地块，只需要统计地块内未遮盖区域的像元，通过部分来代替整体。而全部遮盖的地块，作为缺失值，缺失值的补全可以采用多种方式，包括通过不完全变量自身的插值拟合。以及通过数据集中的完全变量对缺失值进行估计。

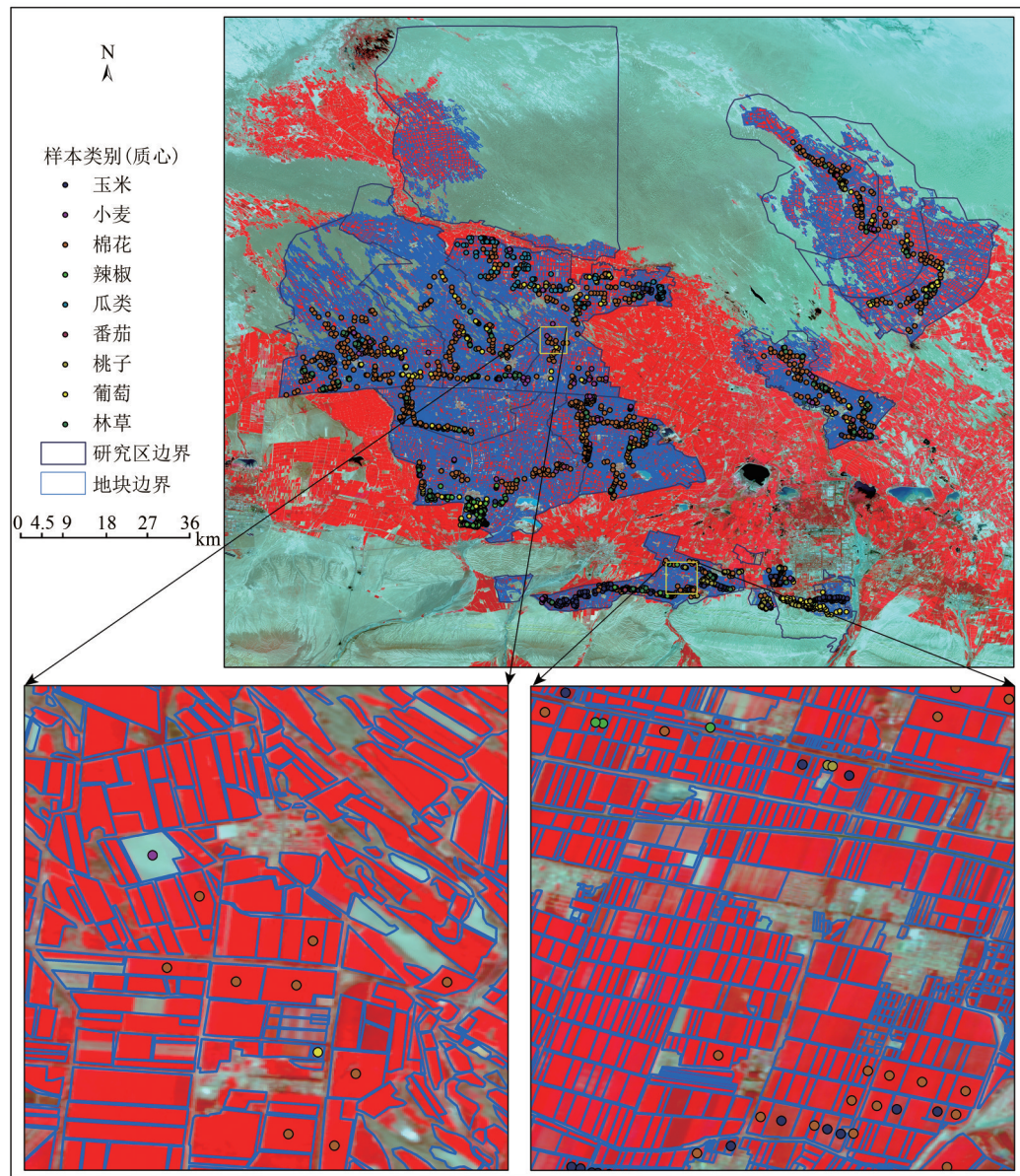


图2 地块数据及样本分布
Fig. 2 Field parcel data and sample distribution

本文使用 Google Earth Engine (GEE) 平台的 USGS Landsat 8 表面反射率数据集。在 GEE 平台利用地块矢量数据统计覆盖研究区地块 4—10 月的可获得全部影像的各波段反射率均值。以数组形式

构建地块上波谱、时相二维特征图作为该地块的二维表征，也是该地块抽象表示。横轴为时相，T0 到 T10 与影像日期顺序相对应。纵轴为波段，band1—band7、band10、band11 共计 9 个波段，与

表面反射率数据集波段构成一致。

表3 Landsat 8像素质量评定波段比特值
Table 3 Landsat 8 pixel quality assessment bit index

位	值	累加和	描述
0	1	1	填充
1	2	3	清晰
2	4	7	水
3	8	15	云阴影
4	16	31	雪
5	32	63	云
云置信度			
6—7	128	255	00=不确定/条件不存在
			01=低(置信度 0—33%)
			10=中(置信度 33%—66%)
			11=高(置信度 66%—100%)
卷云置信度(OLI Band 9获得)			
8—9	256	1023	00=不确定/条件不存在
			01=低(置信度 0—33%)
			10=中(置信度 33%—66%)
			11=高(置信度 66%—100%)
10	1024	2047	地形阻挡

图4(a)所示的地块在T6、T8时相被云覆盖造成区域统计值异常，作为缺失值，这里采用二维 cubic 插值来完成构造造成如图4(b)所示特征图。

3.2 CNN模型及参数设置

在构建CNN的结构时，由于专用架构的多功能性，因此没有标准的过程来搜索超参数和各种类型的层的最佳组合。在这次实验中卷积层、池化层、Dropout 和全连接层相结合，设置的CNN模型构件或参数及其预设值如表4所示。测试了2、3和5的卷积核尺寸和卷积层具有8或16个通道，池化层固定为窗口大小为2的最大池化。Dropout 是一种正则化技术，可在训练过程中将一些神经元随机失活，防止训练较小的数据集时容易造成过拟合提高神经网络的性能。神经元失活的可能性设置为20%、30%和40%。将不同核尺寸的卷积层和最大池化层连接起来以同时处理多尺度特征。利用Flatten层用来将输入“压平”，即把二维的输入一维化，完成从卷积层到全连接层的过渡。每个模型在输出端包含3层的全连接层。最后一层包含9个神经元，对应于分类任务9个类别的概率。全连接各层神经元个数通过考虑最后输出以及Flatten层扁平化的输入来确定的。最后一层分类采用softmax作为激活函数，输出被映射成为(0, 1)的值，而这些值的累和为1（满足概率的性质）。其余各层均采用采用线性整流函数ReLU（Rectified Linear Unit），又称修正线性单元作为激活函数。

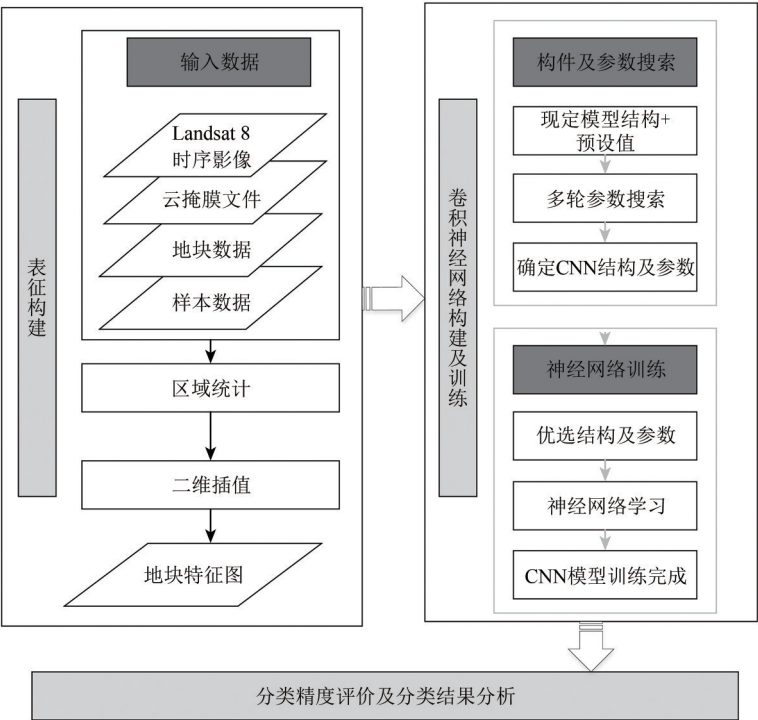


图3 研究技术路线

Fig. 3 Technical route of the study

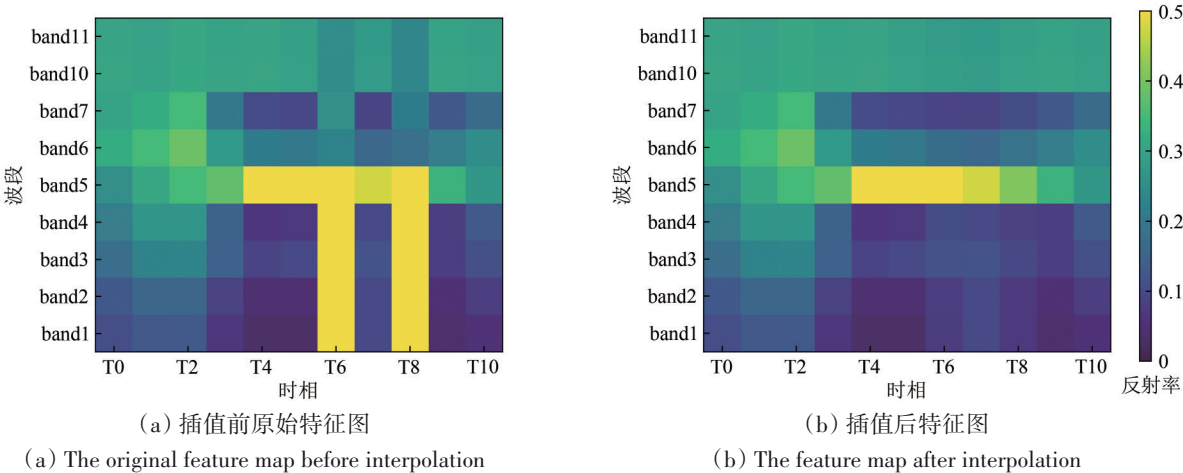


图 4 多时相各波段构成的特征图

Fig. 4 Feature maps of each band composition of multi-temporal RS images

表 4 CNN 模型构件或参数			
Table 4 CNN model components or parameters			
CNN 构件或参数	详细参数	预设值	
卷积层	层数	1、2	
	卷积核	2、3、5	
	通道数	8、16、24	
	激活函数	ReLU	
池化层	窗口大小	2×2	
Dropout	失活概率/%	10、20、30、40	
全连接层	Ⅰ 层	神经元	32
		激活函数	ReLU
	Ⅱ 层	神经元	16
		激活函数	ReLU
	输出层	神经元	9
		激活函数	Sorfmox
优化器	Adam	—	
损失函数	交叉熵损失	—	
训练参数	批量	25、50、75、100	
	批次	每次增加 500 直至达到 停止条件	

在模型的训练过程中，训练次数不是越多越好，对训练集进行过多的学习，会导致过拟合现象，随着学习的进行，对于 CNN 模型权重值学习迭代次数足够多，BP 算法使权值可能收敛过于复杂的决策面，拟合了训练样本中的噪声和其中并不具有代表性的特征。批量训练是将批量样本训练取平均损失来进行梯度计算。批量设置过小，梯度的估值就越不准确，梯度下降的方向波动也越大，难以达到收敛效果。而批次过大，参数的

修正变得缓慢，达到相同的精度所耗损的时间增加。且当批次增大到一定范围内，其确定的下降方向基本不会变化。优化器选择 Adam 进行了训练。遵循原始论文（Kingma 和 Ba，2014）提供的参数作为默认参数（ $\beta_1=0.9$ ， $\beta_2=0.999$ ，学习率为 0.001。）损失函数采用交叉熵损失函数。

3.3 分类精度指标

本文计算了测试集的混淆矩阵、整体精度和 Kappa 系数，以评估所训练分类器的整体性能。计算用户精度、制图精度对单一类别分类性能。总体精度与正确映射的区域成比例，并且适合于区域估计。用户精度表示分类结果中某类别被正确分类的概率，主要用来评价分类结果的可信度。制图精度表示地面的某类别被制图者正确分类的概率。主要用来评价分类方法的好坏。Kappa 系数采用另一种离散多元统计技术，它综合运用了混淆矩阵的所有参数来计算一个最终指标，从而克服了整体精度、用户精度、制图精度指标过分依赖于所选样本点的问题。Kappa 系数的计算公式如下：

$$\text{Kappa} = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})}$$

式中， n 为混淆矩阵中的总列数，也就是类别数； x_{ii} 是混淆矩阵中第 i 行、第 i 列上的样本个数，也就是分类正确的数目； x_{i+} 和 x_{+i} 分别为第 i 行和第 i 列的总样本个数； N 为总样本个数。

4 实验与分析

4.1 训练过程及结果

训练样本占比增加，会使得测试样本占比减少，从而降低总体精度的容错性。这里选用50%作为训练，50%作为验证，这与其他分类器设置一致。在训练CNN时，训练样本的30%留作验证，进行超参数调优设置。即训练集占全体样本的

35%，验证集15%，测试集50%。利用测试集与验证集进行参数设定，搜索过程构建模型的过程是通过测试集的精度来进行多轮的搜索。最初统一采用批量为50，每500次训练后通过测试集分类错误率在两次增加来决定停止一组参数设置下的训练。在对CNN模型的超参数逐步选择完毕，即CNN模型构建完成后开启对训练参数的搜索。其搜索过程如表5所示。

表5 构件或参数搜索过程
Table 5 Component or parameter search process

第1轮搜索												
默认参数	搜索参数	训练批次										
Dropout=0.2; 训练批次=50	卷积层	500	1000	1500	2000	2500	3000	3500	4000	4500	5000	5500
	5×5(8)	9.89	8.52	7.60	7.23	7.88	7.60					
	5×5(16)	9.16	7.42	6.87	6.96	6.87	7.60					
	5×5(24)	9.43	7.42	7.23	7.14	7.97	8.24					
	3×3(8)	10.81	8.15	7.78	7.42	7.23	6.92	6.96	7.60			
	3×3(16)	10.44	7.23	7.51	7.33	6.96	7.05	7.23				
	3×3(24)	8.15	7.05	7.23	7.51							
	2×2(8)	12.91	10.26	9.71	8.42	9.07	9.34					
	2×2(16)	10.07	7.97	8.42	8.15	7.60	7.14	6.96	7.14	7.88		
	2×2(24)	10.07	7.69	7.60	7.42	7.78	7.23	7.05	7.69	6.87	7.14	7.23
	2×2(16)+2×2(16)	12.00	10.44	8.70	8.25	7.69	8.15	8.15				
	2×2(8)+2×2(16)	12.83	8.71	7.42	6.87	6.72	7.42	7.01				
第2轮搜索												
默认参数	搜索参数	训练批次										
卷积层为2×2(8)+2×2(16); 训练批量=50	Dropout	500	1000	1500	2000	2500	3000	3500	4000	4500	5000	5500
	0.1	8.70	7.69	6.78	7.42	6.78	7.14					
	0.2	12.83	8.71	7.42	6.87	6.72	7.42	7.01				
	0.3	11.89	8.42	7.14	6.78	7.51	6.69	7.23	7.03			
	0.4	10.90	9.16	9.43	8.42	8.61	8.33	7.14	7.60	6.96	7.79	7.14
第3轮搜索												
默认参数	搜索参数	训练批次										
卷积层为2×2(8)+2×2(16); Dropout=0.3	批量	500	1000	1500	2000	2500	3000	3500	4000	4500	5000	5500
	25	8.79	7.97	7.60	7.78	6.96	7.33	7.88				
	50	11.89	8.42	7.14	6.78	7.51	6.69	7.23	703			
	75	10.62	8.33	7.78	7.23	6.42	6.71	6.52				
	100	10.71	8.33	7.33	7.68	6.78	6.81	6.91				

注:黑体表示最优结果。

通过前两轮搜索构建的网络结构如图5所示。不能否认的是依然存在大量潜在的网络体系结构并不能全部尝试。更多的是根据经验进行逐

步选择超参数。从相对简单的模型开始，通过更改一个或两个超参数，添加新层，重新排序层或用更复杂的组件替换网络的一部分来生成新模型。

在获取具有良好分类性能的模型被用作启动新一轮搜索的开始。通过这种方式测试模型的大小和

复杂性不断增长，直到分类结果没有进一步改善。最终训练超参数选择为批次为2500，批量为75。

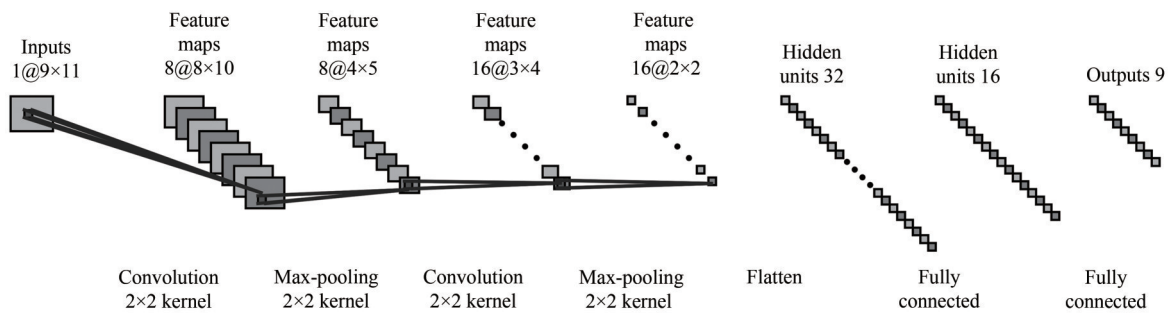


图5 CNN模型结构

Fig. 5 CNN model construction

4.2 分类结果及分析

通过训练的模型对全部的地块表征进行分类，获得如图6所示新疆生产建设兵团第八师2019年的作物分布图。相较于基于像素或者面对对象的分类方法，本算法最大的特点是有着清晰的地块边界，远高于遥感影像的分辨率。

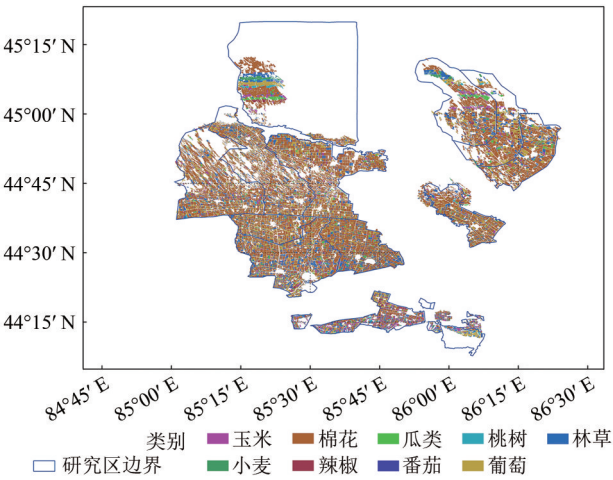


图6 实验区2019年作物分布

Fig. 6 Crop distribution in the experimental area in 2019

通过计算了测试集的混淆矩阵、整体精度和Kappa系数，以评估所训练分类器的整体性能。结果如表3所示。注意到番茄类制图精度都远低于其他类对象，造成该结果的主要原因是番茄地块样本过少，在训练过程中造成对番茄分类的欠拟合，不能对未知的测试样本较好的完成分类任务。棉花作为实验区种植面积最多的作物，其精度高的原因，一方面是样本数量足够，另一方面其收获在10月，生育期与其他作物有着明显区

别。其他类别除了桃树的用户精度为86.96%，林草类的制图精度为82.73%，其余作物的用户与制图精度都高于90%。分类器的整体精度为93.04%、Kappa系数为91.09%。在9类作物分类中获得较好的结果。依据分类器所采用的模型形式可以将监督分类算法分为4类：概率模型、几何模型、逻辑（规则）模型、网络模型。分别选用4类中最为典型机器算作为本算法的对比，包括了基于概率模型的朴素贝叶斯（Naive Bayes）；基于几何模型的支持向量机SVM（Support Vector Machine）；基于逻辑（规则）模型的随机森林RF（Random Forest）以及基于网络模型的多层感知机MLPs（Multilayer Perceptron）。将所获得的分类精度与其他4种常见分类器与本文算法进行对比。为了减少单次分类精度可能由于样本划分带来的偶然误差，对样本集进行10次划分后分别用以上算法进行20次重复实验。结果如图7所示，对比其他算法，可以发现该算法所到的分类模型的平均整体精度为93.91%、平均Kappa系数为92.20%，均高于其他算法。而本文算法受模型训练形式限制，实际上参与训练样本的数量只占全部样本的35%，低于其他算法50%的训练样本，但仍然获得最高的分类精度。分类精度具有饱和性，分类精度越高，提升精度的难度也就越大，分类精度的上限易受到样本质量的限制。可以看出本算法获得了明显的精度提升。其次，20次重复实验也说明了该方法分类结果受到样本划分的精度差异小于其他算法，分类精度在一个较小区间范围，具有更强的稳点性。

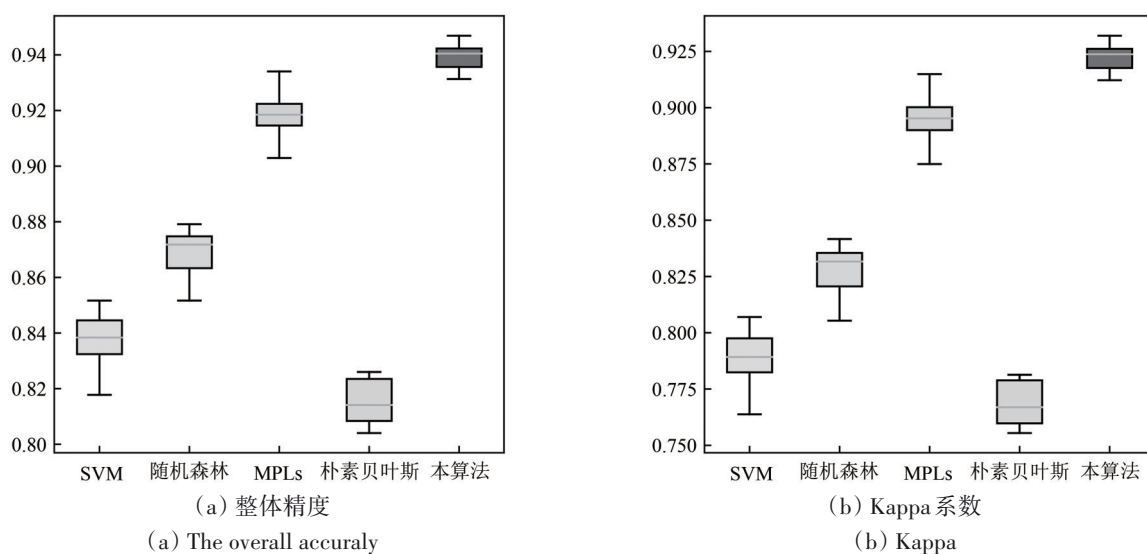


图7 与其他算法的分类精度对比

Fig. 7 Comparison of classification accuracy with other algorithms

5 结 论

本文以新疆生产建设兵团第八师为研究区, 基于全生育期 Landsat 8 影像集与地块数据, 开展作物分类的 CNN 模型的构建。所提出的算法优于传统机器学习分类算法, 是深度学习在农业遥感应用中一次有效的结合。结果表明 CNN 模型在波段、时相二维特征表示中具有巨大潜力, 应被视为基于地块的多时相图像分类任务的优选方法。地块表征的概念将提供分类的信息可以通过整体来看待。表征可以看作为特征的组合或者对特征某种规则的编码。数据组织的方式中也会蕴含着可供分类的信息。最重要的是表征并不仅仅是特征图这一种表现形式。表征作为相对于特征的更高层次表达, 作为联系多类数据获得特定储存形式的深度学习输入。并不需要对遥感数据进行过多选择。这些优势也意味着利用该方法可以减少人类知识及先验假设的加入, 易于实现作物遥感识别的系统集成, 提高分类工作自动化和系统化、流程化程度。

本文研究有着很多的可扩展的方面, 这里列出主要的扩展研究主题以供参考:

(1) 加入其他地块特征。本文在进行构建时相、波谱的特征图作为地块表征时, 使用的是各波段地块地表反射率的均值。这里可以加入其他统计量, 如变异系数、偏度、峰度等反演地块内部反射率综合统计特性。也可以加入纹理分析, 提

取获得地块纹理的定量描述。以此联合构建更高维度的特征图, 通过卷积操作, 获得更能体现地块的综合表征。可以提高分类精度或识别地块间更精细的分类, 如作物品种、地块种植结构等。

(2) 使用多源遥感数据。由于地块数据的存在, 很容易利用地块数据提取地块内影像像元的统计值, 按照波长、影像拍摄时间构成的平面上的点。拥有不同波段设置的多源遥感数据可以很容易的在二维显示。通过二维的插值处理, 获得包含多源数据信息构成的特征图。这张图包含了地块作物生育期内更多时相、波谱信息所构成的复合特征。可以以此开展空间上、时间上的分类任务的泛化。

(3) 加入半监督学习。通过加入更易获得的无标签地块样本来解决样本数量不足问题或改善分类结果。除了传统的半监督学习算法外如标签传播算法; 还包含无监督预训练, 有监督的微调神经网络的训练形式; 以及最近应用广泛的生成对抗网络 (GAN) 的扩展——辅助分类器 GAN (ACGAN), 以生成网络产生伪数据, 利用判别网络对真伪数据进行判断并进行分类, 既使用标签信息进行训练, 同时也重建标签信息。

分类作为机器学习的最主要的任务之一, 诞生了最多的机器学习算法。遥感分类一直得益于机器学习领域的发展。以神经网络为代表深度学习作为机器学习的子集, 近年来发展迅速, 这为遥感分类带来了新的模式。包括农业的各个领域

遥感应用, 都有一个普遍的趋势: 从研究区小范围扩展到大区域; 从数据的单一传感器扩展到多源异构数据共同参与; 从单机模式扩展到高性能并行运算模式。深度学习将与物遥感分类将会有更深入结合, 通过可增长的数据量及运算量来提高作物分类各方面指标, 以完成智能化、系统化、自动化的农作物识别任务。

参考文献(References)

- Anwer R M, Khan F S, Van De Weijer J, Molinier M and Laaksonen J. 2018. Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 138: 74-85 [DOI: 10.1016/j.isprsjprs.2018.01.023]
- Audebert N, Le Saux B and Lefevre S. 2017. Segment-before-detect: vehicle detection and classification through semantic segmentation of aerial images. *Remote Sensing*, 9(4): 368 [DOI: 10.3390/rs9040368]
- Bittner K, Cui S and Reinartz P. 2017. Building extraction from remote sensing data using fully convolutional networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, XLII-1/W1: 481-486 [DOI: 10.5194/isprs-archives-XLII-1-W1-481-2017]
- Blaschke T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1): 2-16 [DOI: 10.1016/j.isprsjprs.2009.06.004]
- Cai Y P, Guan K Y, Peng J, Wang S W, Seifert C, Wardlow B and Li Z. 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sensing of Environment*, 210: 35-47 [DOI: 10.1016/j.rse.2018.02.045]
- Cavallaro G, Riedel M, Richerzhagen M, Benediktsson J A and Plaza A. 2015. On understanding big data impacts in remotely sensed image classification using support vector machine methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(10): 4634-4646 [DOI: 10.1109/JSTARS.2015.2458855]
- Chen Z X, Ren J Q, Tang H J, Shi Y, Leng P, Liu J, Wang L M, Wu W B and Yao Y M. 2016. Progress and perspectives on agricultural remote sensing research and applications in China. *Journal of Remote Sensing*, 20(5): 748-767 (陈仲新, 任建强, 唐华俊, 史云, 冷佩, 刘佳, 王利民, 吴文斌, 姚艳敏. 2016. 农业遥感研究应用进展与展望. *遥感学报*, 20(5): 748-767) [DOI: 10.11834/jrs.20166214]
- Guo R, Liu J B, Li N, Liu S B, Chen F, Cheng B, Duan J B, Li X P and Ma C H. 2018. Pixel-wise classification method for high resolution remote sensing imagery using deep neural networks. *ISPRS International Journal of Geo-Information*, 7(3): 110 [DOI: 10.3390/ijgi7030110]
- Han X B, Zhong Y F and Zhang L P. 2016. Spatial-spectral classification based on the unsupervised convolutional sparse auto-encoder for hyperspectral remote sensing imagery. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-7: 25-31 [DOI: 10.5194/isprs-annals-III-7-25-2016]
- Han Y X. 2018. Study on Per-Field Crop Classification Method Using Remote Sensing Data. Beijing: Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (韩衍欣. 2018. 面向地块的农作物遥感分类方法研究. 北京: 中国科学院遥感与数字地球研究所)
- Han Y X and Meng J H. 2019. A review of per-field crop classification using remote sensing. *Remote Sensing for Land and Resources*, 31(2): 1-9 (韩衍欣, 蒙继华. 2019. 面向地块的农作物遥感分类研究进展. *国土资源遥感*, 31(2): 1-9) [DOI: 10.6046/gtzyyg.2019.02.01]
- Hinton G, Deng L, Yu D, Dahl G E, Mohamed A R, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T N and Kingsbury B. 2012. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6): 82-97 [DOI: 10.1109/MSP.2012.2205597]
- Huang Q T, Qin Z L and Zeng Z K. 2016. Study on the crop classification and planting area estimation at land parcel scale using multi-sources satellite data. *Journal of Geo-Information Science*, 18(5): 708-717 (黄启厅, 覃泽林, 曾志康. 2016. 多星数据协同的地块尺度作物分类与面积估算方法研究. *地球信息科学学报*, 18(5): 708-717) [DOI: 10.3724/SP.J.1047.2016.00708]
- Kingma D P and Ba J. 2014. Adam: a method for stochastic optimization//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego: [s.n.]
- LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. *Nature*, 521(7553): 436-444 [DOI: 10.1038/nature14539]
- Li H Y, Su A Y, Liu C, Wu Y Q and Chen S B. 2019. Bisupervised network with pyramid pooling module for land cover classification of satellite remote sensing imagery. *Journal of Applied Remote Sensing*, 13(4): 048502 [DOI: 10.1117/1.JRS.13.048502]
- Li Z Y, Wu B F, Zhang M, Xing Q, Li M Y and Yan N N. 2019. Identifying rapeseed planting area using an object-oriented method and crop phenology. *Journal of Geo-Information Science*, 21(5): 720-730 (李中元, 吴炳方, 张淼, 邢强, 李名勇, 闫娜娜. 2019. 利用物候差异与面向对象决策树提取油菜种植面积. *地球信息科学学报*, 21(5): 720-730) [DOI: 10.12082/dqxxkx.2019.180345]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Lu C, Yang X M, Wang Z H and Li Z. 2018. Using multi-level fusion of local features for land-use scene classification with high spatial resolution images in urban coastal zones. *International Journal of Applied Earth Observation and Geoinformation*, 70: 1-12 [DOI: 10.1016/j.jag.2018.03.010]
- Onojeghuo A O, Blackburn G A, Huang J F, Kindred D and Huang W J. 2018. Applications of satellite 'hyper-sensing' in Chinese agri-

- culture: challenges and opportunities. *International Journal of Applied Earth Observation and Geoinformation*, 64: 62-86 [DOI: 10.1016/j.jag.2017.09.005]
- Ping Y P. 2016. Crop Classification Based On Analysis of Phenological Characteristics of MODIS Times Serise. Harbin: Harbin Normal University (平跃鹏. 2016. 基于MODIS时间序列地表物候特征分析及农作物分类. 哈尔滨: 哈尔滨师范大学) [DOI: 10.7666/d.D831058]
- Rahman M R and Saha S K. 2008. Multi-resolution segmentation for object-based classification and accuracy assessment of land use/land cover classification using remotely sensed data. *Journal of the Indian Society of Remote Sensing*, 36(2): 189-201 [DOI: 10.1007/s12524-008-0020-4]
- Seal A, Garcia-Pedrero A, Bhattacharjee D, Nasipuri M, Lillo-Saavedra M, Menasalvas E and Gonzalo-Martin C. 2020. Multi-scale RoIs selection for classifying multi-spectral images. *Multidimensional Systems and Signal Processing*, 31(2): 745-769 [DOI: 10.1007/s11045-019-00684-1]
- Shao Z F, Yang K and Zhou W X. 2018. Performance evaluation of single-label and multi-label remote sensing image retrieval using a dense labeling dataset. *Remote Sensing*, 10(6): 964 [DOI: 10.3390/rs10060964]
- Shi Z, Liang Z Z, Yang Y Y and Guo Y. 2015. Status and Prospect of agricultural remote sensing. *Transactions of the Chinese Society for Agricultural Machinery*, 46(2): 247-260 (史舟, 梁宗正, 杨媛媛, 郭燕. 2015. 农业遥感研究现状与展望. 农业机械学报, 46(2): 247-260)
- Shin H C, Roth H R, Gao M C, Lu L, Xu Z Y, Nogues I, Yao J H, Molura D and Summers R M. 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5): 1285-1298 [DOI: 10.1109/TMI.2016.2528162]
- Su C, Feng C J, Huang Z C and Zhang X C. 2018. Extraction of rice cropping area from high resolution remote sensing image based on sample knowledge mining//*Proceedings of the 2018 7th International Conference on Agro-Geoinformatics*. Hangzhou: IEEE: 1-7 [DOI: 10.1109/Agro-Geoinformatics.2018.8475988]
- Sun C L, Bian Y, Zhou T and Pan J J. 2019. Using of multi-source and multi-temporal remote sensing data improves crop-type mapping in the subtropical agriculture region. *Sensors*, 19(10): 2401 [DOI: 10.3390/s19102401]
- Valero S, Morin D, Inglada J, Sepulcre G, Arias M, Hagolle O, Dedieu G, Bontemps S, Defourny P and Koetz B. 2016. Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. *Remote Sensing*, 8(1): 55 [DOI: 10.3390/rs8010055]
- Wang H Z, Wang Y, Zhang Q, Xiang S M and Pan C H. 2017. Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote Sensing*, 9(5): 446 [DOI: 10.3390/rs9050446]
- Wei W, Zhang J Y, Zhang L, Tian C N and Zhang Y N. 2018. Deep cube-pair network for Hyperspectral imagery classification. *Remote Sensing*, 10(5): 783 [DOI: 10.3390/rs10050783]
- Zabalza J, Ren J C, Zheng J B, Zhao H M, Qing C, Yang Z J, Du P J and Marshall S. 2016. Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing*, 185: 1-10 [DOI: 10.1016/j.neucom.2015.11.044]
- Zhang C, Pan X, Li H P, Gardiner A, Sargent I, Hare J and Atkinson P M. 2018. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140: 133-144 [DOI: 10.1016/j.isprsjprs.2017.07.014]
- Zhang H X, Li Q Z, Liu J G, Shang J L, Du X, McNairn H, Champagne C, Dong T F and Liu M X. 2017. Image classification using RapidEye data: integration of spectral and textual features in a random forest classifier. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(12): 5334-5349 [DOI: 10.1109/JSTARS.2017.2774807]
- Zhang L F, Zhang L P, Tao D C and Huang X. 2012. On combining multiple features for Hyperspectral remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3): 879-893 [DOI: 10.1109/TGRS.2011.2162339]
- Zhong L H, Hu L N and Zhou H. 2019. Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, 221: 430-443 [DOI: 10.1016/j.rse.2018.11.032]

Crop classification based on two-dimensional representation and CNN model from remote sensing

ZHANG Qiankun^{1,2}, MENG Jihua², REN Chao¹

1. College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China;

2. Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

Abstract: Crop planting area and its spatiotemporal distribution information are crucial for agricultural management, structural adjustment of the planting industry, national food security, and other fields. Remote sensing can extract crop planting information quickly. However, a

large amount of ground survey information, expert knowledge, and manual correction operation after classification are needed to meet actual production needs. Using remote sensing and existing geographic information data for intelligent information extraction is the future development trend. The purpose of this study is to investigate the classification representation method of the deep learning model on the basis of the constraints of field parcel data. Multi-source heterogeneous data are linked by constructing representation. The method can identify the spectral differences of different crops in different phases and classify the crop types.

The study area is in the eighth division of Xinjiang Production and Construction Corp. The platform of Google Earth is used to obtain all Landsat 8 images during the crop growth period of the research area in 2019. The image data and their quality assessment band on GEE are utilized to determine the statistical or estimated values of phases in the study area. Then, the average reflectivity of each band of the extracted block level is arranged in accordance with the time-phase sequence and wavelength, and a plot representation of the spectrum and time-phase two-dimensional properties is constructed. The construction of land plot representation realizes the connection of geographic information and remote sensing data and makes the application of deep learning in crop remote sensing classification possible. The completed plot representation is used to construct and train a Convolution Neural Network (CNN) model.

A step-by-step optimization process is implemented to search for the best combination of super parameters and various types of layers. Construction of a relatively optimal CNN model is obtained, and classification of the constructed feature map is carried out based on this model to complete the classification of the research area in 2019. We obtain a crop distribution map whose resolution is much higher than that of the remote sensing image. Moreover, the overall accuracy of the CNN model reaches 93.04%, and the kappa coefficient is 91.09%. The results of nine kinds of crop classification are good. After thousands of rounds of data learning, the proposed method exhibits lower classification error fluctuation and higher stability than other machine learning algorithms.

The research object is plot representation, which is the abstract expression of plot planting information. It can be used as the standard input of the deep learning model. Through the construction of plot representation, crop classification can be indirectly identified by remote sensing. The proposed method has outstanding application potential in land feature representation and crop classification and should be regarded as an optimal method for multi-temporal image classification tasks based on field parcel data. The method can be used as a reference in the application of deep learning in remote sensing.

Key words: remote sensing, crop classification, CNN, feature map, plot representation, multi-temporal image, per-field