

# 迁移深度卷积神经网络模型秋粮作物泛化识别

张凤<sup>1,2,3</sup>, 张锦水<sup>1,2,3,4</sup>, 段雅鸣<sup>1,2,3</sup>, 杨志<sup>1,2,3</sup>

1. 北京师范大学 遥感科学国家重点实验室, 北京 100875;

2. 北京师范大学 地理科学学部 北京市陆表遥感数据产品工程技术研究中心, 北京 100875;

3. 北京师范大学 地理科学学部 遥感科学与工程研究院, 北京 100875;

4. 青海师范大学 高原科学与可持续发展研究院, 西宁 810016

**摘要:** 深度卷积神经网络DCNN (Deep Convolutional Neural Networks) 具有“端到端”、高效提取多尺度语义特征的优点, 被广泛应用于遥感农作物识别中, 取得了重要的进展。但深度学习模型需要大量标签样本, 获取费时费力、成本高, 限制了其广泛使用。本文基于模型迁移学习策略, 将美国农业部统计署作物数据层CDL (Cropland Data Layer) 作为作物标签数据, 利用伊利诺伊州和印第安州的作物生长季 Landsat OLI 卫星影像训练2016年—2019年4个单年份模型和2016年—2017年、2016年—2018年两个多年份的U-net模型, 将模型直接在美国3个测试区和中国黑龙江省黑河市的2016年—2020年迁移泛化分类。研究结果表明: 1) 基于CDL训练的U-net模型能够在美国本土推广应用, 美国3个测试区的2016年—2020年的时间泛化总体精度在80%以上, 玉米识别精度高于大豆, 说明了模型在空间上具有很好的迁移能力。(2) 对于黑河市秋粮, 多年玉米识别总体精度平均高出大豆3%, 原因是玉米种植地块更为规整、种植规模更大, 单年份秋粮识别总体精度69%—79%, 但多年份模型要优于单年份模型, 这可能是随标签样本数量增多提升了训练样本的代表性, 将中国与美国秋粮种植的差异通过样本数量的增加得以弥补, 模型迁移到中国黑河区域的精度均低于美国本土, 这是由于洲际间气候、作物物候、种植农业景观等差异导致遥感特征不一致, 导致模型泛化能力降低。

**关键词:** 遥感, 迁移学习, CDL, 时空泛化, 大豆, 玉米

**中图分类号:** P236/P2

**引用格式:** 张凤, 张锦水, 段雅鸣, 杨志. 2024. 迁移深度卷积神经网络模型秋粮作物泛化识别. 遥感学报, 28(3): 661–676

Zhang F, Zhang J S, Duan Y M and Yang Z. 2024. Transferring deep convolutional neural network models for generalization mapping of autumn crops. National Remote Sensing Bulletin, 28(3): 661–676 [DOI: 10.11834/jrs.20241360]

## 1 引言

准确、快速地获取秋粮作物种植空间分布和面积对于粮食安全和地区农业政策制定意义重大(谢登峰等, 2015)。遥感具有覆盖范围大、视点高、现势性强的优势, 是当前进行农作物提取的最重要技术手段之一(胡琼等, 2015)。因此, 利用遥感技术进行作物面积测量一直是农业遥感的热点领域。

自对地观测遥感技术发展以来, 国内外专家学者注重发展模式识别的模型方法进行秋粮作物类型识别(刘佳等, 2018; 许淇等, 2019; Yang等,

2011; 马丽等, 2008), 也有研究利用多源遥感特征进行作物识别(Zhang等, 2018; Phalke等, 2020; You和Dong, 2020)。然而, 迄今为止秋粮作物遥感分类依然存在一些不足, 主要体现在传统分类方法多利用较为单一的遥感响应特征, 无法提取深层次特征, 分类模型能力不足, 这种人为选择特征训练出的模型一般要求训练区域和测试区域遥感特征要严格一致, 因此一般模型适应特定区域, 难以大范围推广应用(Zhong等, 2020)。

深度学习通过学习浅层细节特征, 随着模型层数的加深不断提取更为复杂的特征, 在不需要人为干预的情况下不断改进预测, 在很大程度上能

收稿日期: 2021-05-31; 预印本: 2022-10-19

基金项目: 国家自然科学基金重大项目(编号:42192580, 42192584); 高分辨率对地观测系统重大专项(民用部分)(编号:20-Y30F10-9001-20/22)

第一作者简介: 张凤, 研究方向为农业遥感。E-mail: 201821051194@mail.bnu.edu.cn

通信作者简介: 张锦水, 研究方向为农业遥感。E-mail: zhangjs@bnu.edu.cn

够克服传统分类器的不足,自动、高效地提取出大量的深层特征,实现端到端的自动化分类,已被广泛引入到遥感应用中(宋德娟,2019;Jiang等,2018;Zhong等,2019),在农作物遥感分类中也得到了广泛应用(Zhong等,2020;Sun等,2020)。然而,深度卷积神经网络模型准确分类的前提是需要大量高质量的作物类型标签数据,但这种作物标签数据的获取需要耗费大量的人力物力(Xu等,2020;许晴等,2022),这是深度学习面向实际应用最大的瓶颈之一。

美国农业部统计署(National Agricultural Statistics Service)每年定期发布基于CLU(Common Land Unit)和时间序列遥感数据生产的作物类型数据CDL(Cropland Data Layer)([https://www.nass.usda.gov/Research\\_and\\_Science/Cropland/SARS1a.php](https://www.nass.usda.gov/Research_and_Science/Cropland/SARS1a.php) [2021-05-31],主要农作物的制图精度在85%—95%(Egli,2008)。已有相关研究将CDL数据作为标签数据构建深度学习模型进行作物分类,取得了理想的结果,如Cai等(2018)利用CDL作为标签样本进行县级尺度的模型训练,测试了多个年份、不同区域的模型迁移应用,验证了该分类模式的适用性;Xu等(2020)利用Landsat ARD(Landsat Analysis Ready Data)和CDL数据构建训练数据集构建DCM(Deep Crop Mapping)模型实现了位于美国6个实验站点玉米、大豆的空间泛化实验,试验站点平均Kappa为85.8%,泛化区域的平均Kappa为82.0%。Konduri等(2020)采用可扩展的聚类—标签模型,利用多年MODIS NDVI和CDL构建数据集进行训练,构建相似物候区域,基于物候区和作物类别之间的空间一致性,利用作物曲线将作物标签类型分配到每一个物候区,在玉米、大豆和冬小麦的主产区,8种主要作物分类精度达到了70%。

已有研究证实了CDL能够作为作物标签数据构建小区域分类模型应用到其他区域或年份,本文利用CDL构建标签样本进行深度学习模型训练,验证模型在更大空间和时间尺度上的迁移泛化能力,提高秋粮遥感识别效率。本研究以美国的CDL作为标签数据训练深度学习模型,迁移模型到美国其他区域和中国东北地区,验证模型在不同农业景观、不同年份的时空迁移能力,解释CDL作为标签数据训练模型在大范围进行农作物分类的适用性,为大尺度农作物自动、高效的分类提供理论依据。

## 2 研究区与数据

### 2.1 研究区概况

本研究选择美国中西部玉米、大豆主产区的CDL作为训练样本。为验证模型跨越时间、空间尺度上的泛化能力,选取美国玉米和大豆种植较为集中的印第安纳州和伊利诺伊州作为训练区,跨度范围 $36^{\circ}58'N$ — $42^{\circ}30'N$ 、 $84^{\circ}46'W$ — $94^{\circ}31'W$ ,共约24万 $km^2$ ,玉米和大豆占到作物种植面积的55%以上,位于平原地区,田块比较规则,面积大小约33.33—66.67 $hm^2$ ;模型迁移泛化区是美国大豆、玉米种植比较集中的区域,分别位于明尼苏达州南部、爱荷华州中西部和俄亥俄州东北部(分别记为A1、A2、A3,范围分别为 $44^{\circ}35'N$ — $44^{\circ}40'N$ 、 $99^{\circ}47'W$ — $99^{\circ}52'W$ 、 $41^{\circ}50'N$ — $41^{\circ}55'N$ 、 $96^{\circ}32'W$ — $96^{\circ}37'W$ 、 $40^{\circ}22'N$ — $40^{\circ}27'N$ 和 $83^{\circ}02'W$ — $83^{\circ}07'W$ ),每个区域平均面积为3500 $km^2$ ,测试区中心点之间的南北纬度相差 $3^{\circ}$ ,东西经度相差 $10^{\circ}$ 。

为研究洲际跨度的模型迁移能力,本文选择了中国黑龙江省黑河市( $47^{\circ}42'N$ — $51^{\circ}03'N$ , $124^{\circ}45'E$ — $129^{\circ}18'E$ ),该地区地处嫩江平原的东部,大兴安岭的东端,境内水系发达,东部为山区,西部耕地较集中,山地交错分布;气候类型属于寒温带大陆性季风气候,春夏湿润,适宜种植玉米、大豆,该地区的田块相对规则,大小平均在1—1.6 $hm^2$ ,玉米和大豆种植面积之和达到了95%以上,二者比例约为3:7。本文利用CDL训练的模型迁移黑河市进行作物提取,实现空间尺度的泛化,目的是测试秋粮深度学习分类模型的空间泛化能力,从更大地域空间尺度分析不同农业景观和不同年份的时空迁移能力,验证基于CDL训练深度学习模型的泛化能力。

### 2.2 数据集及处理

#### 2.2.1 训练数据集构建

本研究所采用的遥影像数据是2016年—2020年作物生长关键物候期(播种期和生长旺盛期)的Landsat 8 OLI地表反射率,空间分辨率为30 m,波段选取蓝、绿、红、近红外、短波红外1和短波红外2(Xu等,2020)。GEE(Google Earth Engine)平台提供了大量存档的Landsat影像和CDL数据。本

文选择作物的播种期和生长旺盛期两个时期作为识别关键期，考虑到遥感影像受到云的影响，需要对一定时间窗口内的遥感影像进行晴空影像合成处理，具体流程见图 1：首先利用 GEE 提供的 Fmask 算法计算出云分数，使用 QA 波段和按位与运算逐像元进行中值合成去云处理；然后将影像统一转成 UTM 投影，WGS-84 地理坐标系。从美国农业部网站获取的农业数据分析可以得到，美国农业种植区域广，作物种植物候区域间差异比较大，本文在大区域尺度上选择播种和生长旺盛的遥感影像分别为 3 月 15 日至 6 月 15 日和 7 月 1 日至 10 月 1 日时间段，通过两个时期影像合成晴空影像进行模型训练。表 1 是训练遥感影像选取时间窗口和目标作物（玉米、大豆）种植结构。

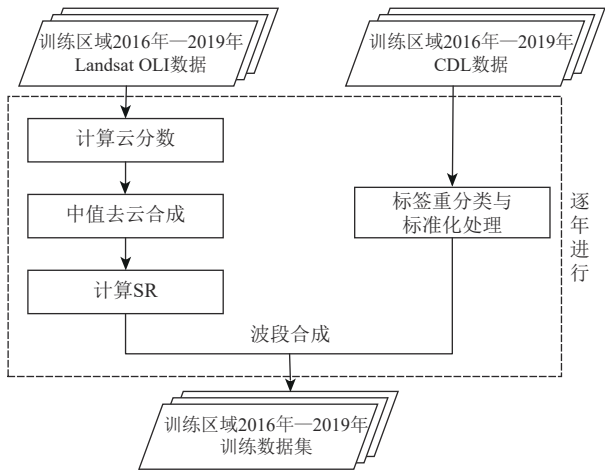


图1 训练数据集构建流程

Fig. 1 The flowchart of the training data set construction

图 1 是训练数据集构建技术流程，选取训练区域 2016 年—2019 年秋粮作物生长季的影像，进行

去云处理，选取训练区域 2016 年—2019 年 CDL 数据作为标签数据，在 GEE 平台进行重分类和影像标准化，即将原有的作物标签重分类为模型训练需要的自然数序列（0-玉米；1-大豆；2-其他；3-背景值），并与训练区域的遥感影像波段叠加合成为训练数据集。

表 1 伊利诺伊州和印第安纳州训练数据集遥感影像及作物样本信息

Table 1 Detailed information of the remote sensing image data of training dataset and samples of Illinois and Indiana

年份	时段	玉米面积占比/%	大豆面积占比/%
2016	03-15—06-15	30.83	25.70
	07-01—10-01		
2017	03-15—06-15	28.06	27.82
	07-01—10-01		
2018	03-15—06-15	27.13	27.60
	07-01—10-01		
2019	03-01—06-15	27.27	26.00
	07-01—10-01		

2.2.2 测试数据集构建

本研究为分析秋粮分类模型的时空尺度上的泛化能力，选取了美国本土 3 个农作物集中种植区和中国黑河市作为测试区域，对 2016 年—2020 年秋粮提取进行时间泛化研究。分析黑河市农业统计资料，该地区秋粮物候种植与美国研究差异比较大，选择 4 月 1 日至 6 月 20 日、7 月 1 日至 10 月 15 日 2 个时间段与训练区时间对应。

表 2 测试区遥感影像选取时间详见表 2，为保证测试区域作物生长季影像的完备，本研究适当调整了影像选取的时间窗口。预测数据集的预处理流程与训练数据集影像处理部分一致。

表 2 测试区域遥感影像选取时间段

Table 2 Time window for remote sensing image selection in test area

年份	时间	时间段			
		区域 A1	区域 A2	区域 A3	黑河市
2016	第一期	03-15—06-15	03-15—06-15	03-15—06-15	04-01—06-20
	第二期	07-01—10-15	07-01—10-15	07-01—10-15	07-01—10-15
2017	第一期	03-15—06-15	03-15—06-15	03-01—06-15	04-01—06-20
	第二期	07-01—10-15	07-01—10-15	07-01—10-15	07-01—10-01
2018	第一期	03-15—06-15	03-15—06-15	03-15—06-15	04-01—06-20
	第二期	07-01—10-15	07-01—10-15	07-01—10-15	07-01—10-01
2019	第一期	03-15—06-15	03-15—06-15	03-15—06-15	04-01—06-20
	第二期	07-15—10-15	07-01—10-15	07-01—10-15	07-01—10-15
2020	第一期	03-01—06-15	03-15—06-15	03-15—06-15	04-01—06-20
	第二期	07-01—10-15	07-01—10-15	07-01—10-15	07-01—10-01



2.2.3 精度评价数据

对于美国3个测试区域（A1、A2、A3），本研究采用2016年—2020年的CDL数据作为参考数据

进行精度评价，黑河市是使用2016年—2020年的春播作物野外实测数据进行精度评价。图2是各测试区参考数据分布图。

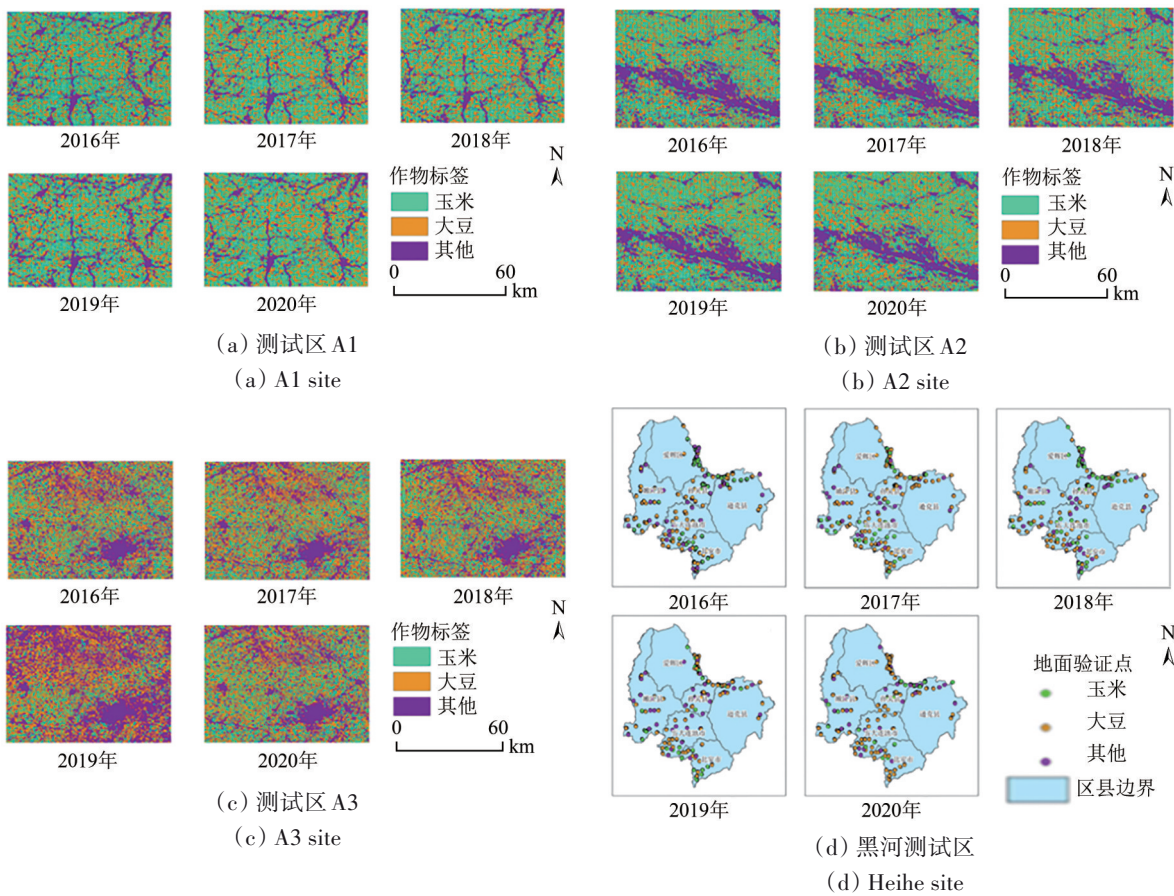


图2 各测试区参考数据  
Fig. 2 Reference data of all test sites

表3 黑河市2016年—2020年春播作物野外样点个数  
Table 3 Number of field samples of spring sown crops in Heihe City from 2016 to 2020

年份	玉米	大豆	其他
2016	1656	2474	2657
2017	533	1620	1232
2018	953	1428	1112
2019	642	1718	1045
2020	725	2665	1284

3 实验方法

3.1 U-net分类模型

近年来，深度全卷积网络FCN（Fully Convolutional Network）因具有学习高层次特征的能力，在语义分割领域取得长足的进展。FCN是

一种端到端深度网络结构，通过卷积层下采样扩展感受野，增加上下文信息，提高分类精度；通过添加上采样层，使输出图像和输入图像的尺寸保持一致，实现逐像素分类。U-net是继承FCN优点的深度卷积神经网络（Ronneberger等，2015），该模型是U型对称网络，左右两侧分别对应编码和解码，编码结构主要作用是降低维度、提取图像特征，解码结构主要作用是恢复图像细节特征和空间、维度信息；解码和编码的对应层相连接，恢复提取目标的细节信息和图像分辨率。近些年来已有研究将U-net模型应用于多波段遥感数据分类，Pan等（2020）基于U-net深度学习架构，利用Worldview影像对广州市复杂城中村的单个建筑进行分割和分类，U-net模型的建筑分割整体精度达到86%以上，分类整体精度达到83%以上，Wei



等 (2019) 将归一化方法引入到 U-net 模型中, 利用 Sentinel-1 时空数据进行 2017 年吉林省扶余市农作物制图, 总体精度达到 85%, Kappa 系数为 0.82。作为典型的 FCN 网络之一, U-net 模型应用于多波段遥感数据分类, 可以克服样本量小、样本量不平衡问题 (Ronneberger 等, 2015)。

本文基于迁移学习策略的 U-net 架构在编码阶段共有 6 层, 输入图像大小从 256 逐渐被压缩至 4 像素, 本文的网络架构如图 3 所示。在解码阶段, 大部分上采样方法均采用反卷积方式, 在输出层前所用的上采样方法为双线性内插法。同时, 为保证提升模型训练速度, 所有特征提取模块均采用残差卷积模块 (Ronneberger 等, 2015)。

### 3.2 实验设计

本研究共设计 6 个实验场景开始模型的训练, 见图 3, 分别用 2016 年、2017 年、2018 年、2019 年美国的 CDL 作为标签数据进行模型训练, 分别记为 model\_2016、model\_2017、model\_2018 和 model\_2019, 为分析在训练数据集中加入多年份训练数据是否能提升模型的泛化能力, 使用 2016 年和 2017 年的数据训练了 model\_16\_17, 使用 2016 年—2018 年的数据训练了 model\_16\_17\_18, 共生成 6 个秋粮识别模型。

为对比不同模型迁移泛化性能, 本研究训练了一个单年份的随机森林模型进行模型迁移作物时空泛化分类, 对比分析传统机器学习模型和深度学习模型之间的时空泛化差异。

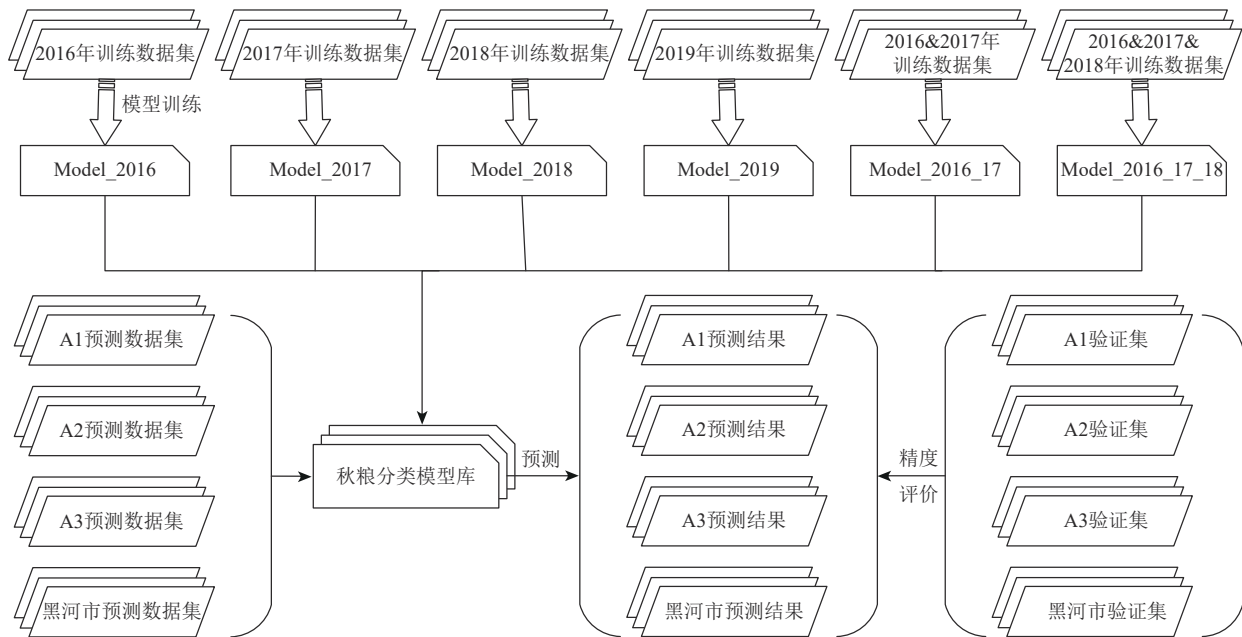


图3 实验设计思路

Fig. 3 Experimental design idea

#### 3.2.1 模型训练与预测

本研究使用深度学习框架是卷积神经网络框架 Caffe (Convolutional architecture for fast feature)。模型超参数设置参考了 Wei 等 (2019), 分别为初始学习率为  $10^{-4}$ , 下降策略为 inv, 学习率变化指数为 0.0001, power 为 0.75, 梯度下降算法设置为 Adam, 权值衰减为 0.0005; 单年份训练数据的 batch\_size 为 32, model\_2016\_17 的 batch\_size 为 64, model\_2016\_17\_18 的 batch\_size 为 128。本研究训练数据集和预测数据均被切分为  $256 \times 256$  个像素大

小斑块, 相邻斑块之间重叠度为 128 个像素。本文采用交叉信息熵构建损失函数。

$$\text{Loss} = - \sum_{i=1}^C \left\{ (y == i) \times \log \frac{e^{a_i}}{\sum_{i=1}^C e^{a_i}} \right\} \quad (1)$$

式中,  $C$  为总类别数, 本研究  $C=4$ , 分别对应玉米、大豆、其他和背景值, 则  $i=0, 1, 2, 3$ ;  $y$  是对应的真值,  $(y==i)$  为关系运算, 若  $y=i$ , 值为 1, 若  $y \neq i$ , 值为 0;  $a_i$  是模型最后一层全卷积层输出对应真值类别  $i$  的概率。

在模型训练阶段, 迁移学习预训练U-net网络, 使用训练数据集对模型进行训练, 模型最后通过全卷积层与 SoftmaxWithLoss 层相连接计算训练 Loss, 进行梯度下降、权值更新保证 Loss 的收敛, 直接 Loss 收敛满足阈值要求确定最终训练模型。

在预测阶段, 将U-net模型中的 SoftmaxWithLoss 层改为 Softmax 层, 输出结果每一层对应每种地物类别的归属概率值, 每个像元的最终类别由对应概率最高的类别所确定, 最终生成作物专题图。

### 3.2.2 精度评价方法

本研究采用混淆矩阵定量评价识别模型的精度, 采用总体精度 OA (Overall Accuracy)、F1 分数 (F1-score) 作为精度评价指标。

(1) 总体精度: 分类问题中所有像元被正确分类的像元数占像元总数的比值, 其计算公式如下:

$$P_o = \frac{\sum_{k=1}^n p_{kk}}{p} \quad (2)$$

式中,  $P_o$  表示总体精度,  $p_{kk}$  表示第  $k$  类中被分类正确的像元数,  $n$  为总类别数,  $p$  为总样本数。

(2) F1-分数: F1 分数综合考虑某一类别的制图

精度与用户精度影响, 是一种更为均衡、综合的分类衡量指标 (赵英时, 2003)。计算公式如下所示:

$$F1 = \frac{2 \times (P_u \times P_p)}{P_u + P_p} \quad (3)$$

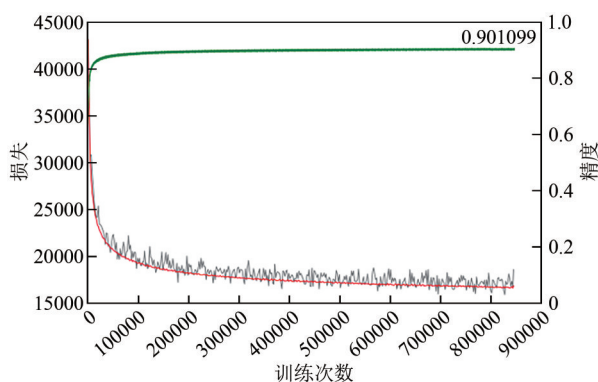
式中,  $P_u$  是用户精度, 即被分为某类的像元中正确分类的像元数与该类像元总数的比值;  $P_p$  表示制图精度, 即被分为某类的像元中正确分类的像元数与真值中该类像元总数的比值。

## 4 结果与分析

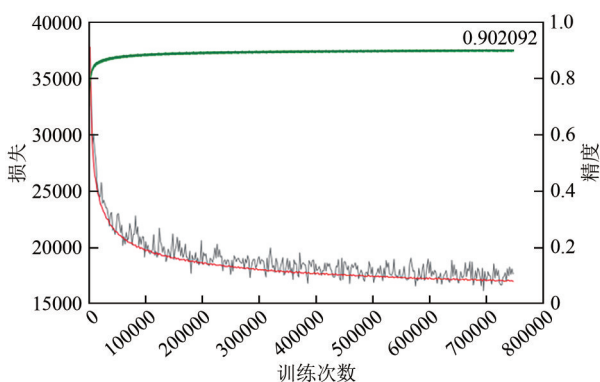
### 4.1 模型收敛能力

根据3.2节中的方法和超参数设置, 本研究训练6个分类模型, 不同模型的收敛速度如图4所示。

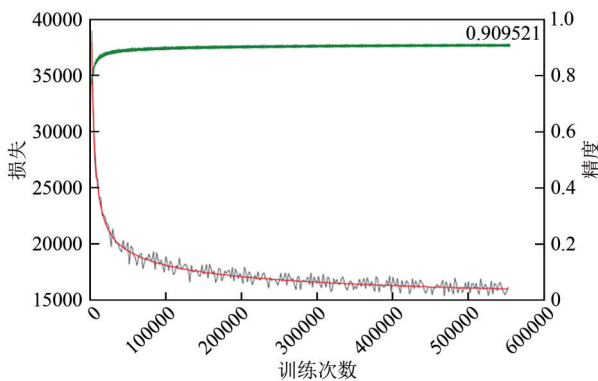
图4中, 前4个 loss 图是单年份数据训练模型 loss 和验证精度的变化过程, 后两个 loss 图是两个多年份数据训练模型的 loss 变化图。在模型训练过程, 4个单年份模型的迭代次数均到了50万次, 模型验证精度维持在90%以上; 由于 model\_2016\_17 和 model\_2016\_17\_18 对应两个年份模型的数据量增大, 导致训练速度下降, 耗时更长, 验证精度也达到了90%以上。



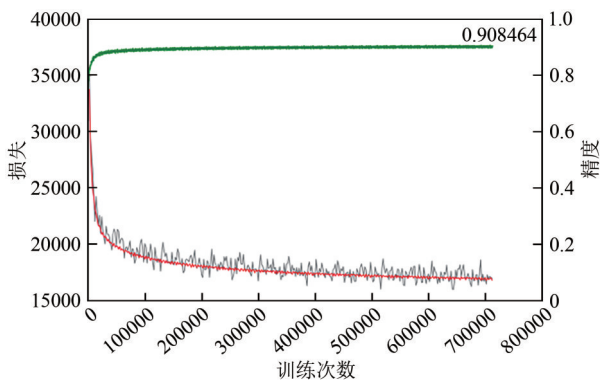
(a) Model\_2016



(b) Model\_2017



(c) Model\_2018



(d) Model\_2019

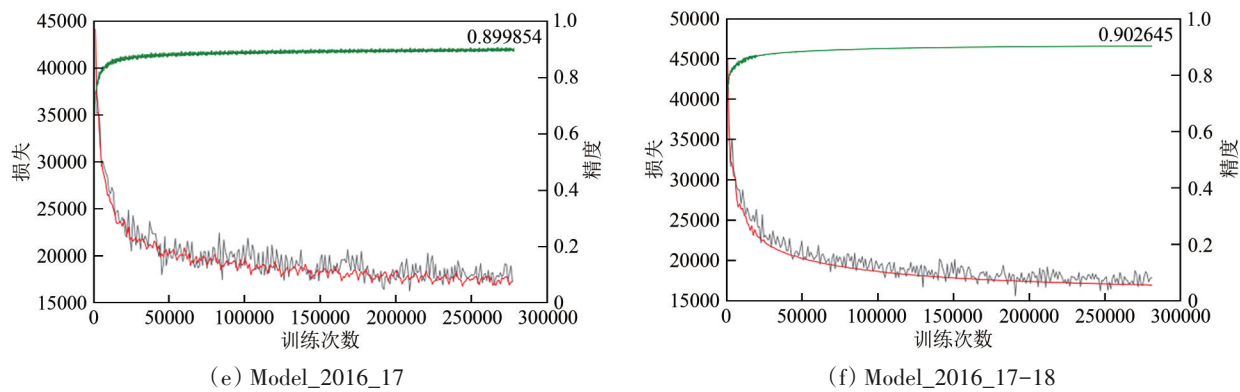


图4 模型训练 Loss 图  
Fig. 4 Loss diagram of model training

4.2 作物识别结果

图5是model\_2016在美国3个测试区2016年—2020年的玉米、大豆预测结果；图6是model\_2016\_17\_18和model\_2016在黑河市测试区的玉米、大豆预测结果。从两个地区预测结果发现，在美

国测试区能够准确地识别玉米和大豆，地块较完整，边缘清晰；但在黑河测试区，单年份模型多个年份玉米和大豆均存在漏分的现象，地块形状表现的不完整。

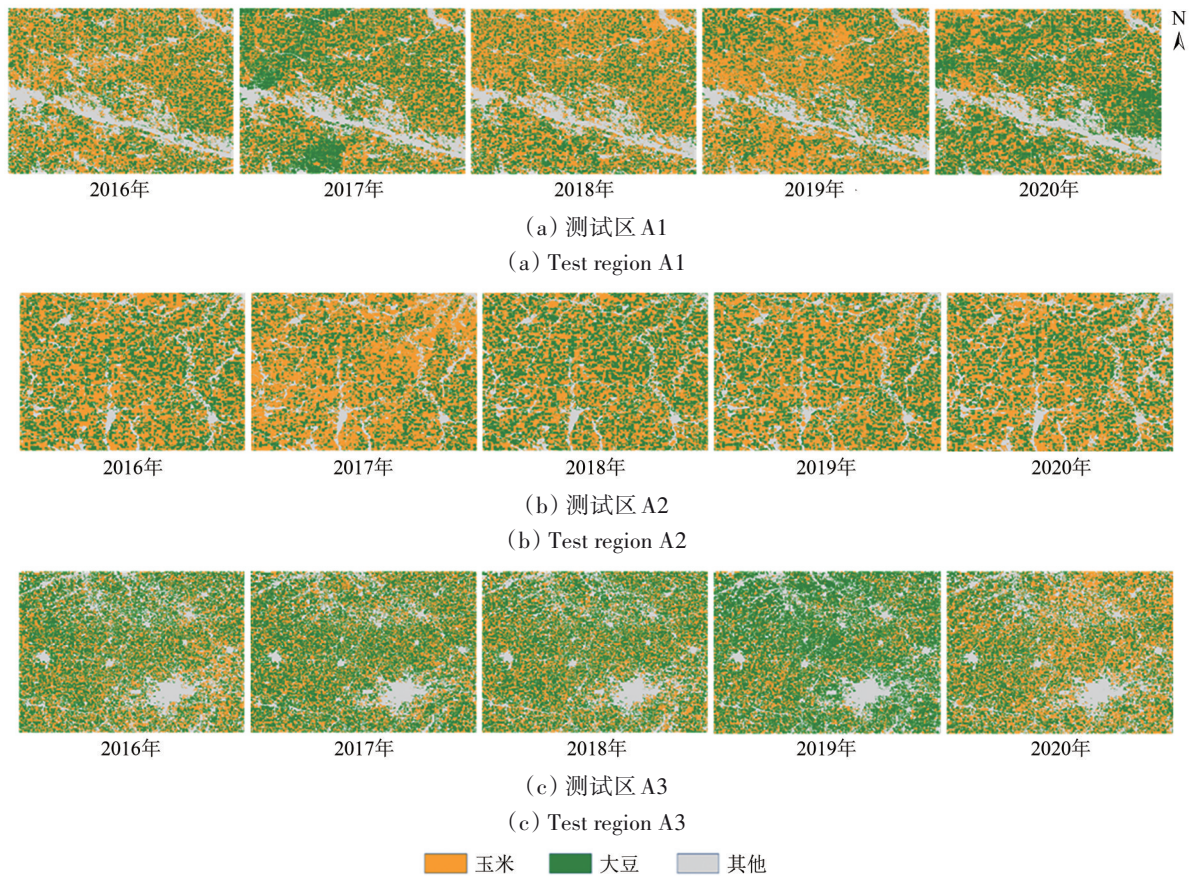


图5 model\_2016在美国3个测试区玉米和大豆的预测结果  
Fig. 5 Prediction results of maize and soybean in US test area by model\_2016

图7和表4是单年份数据训练的4个模型分别在测试区域预测当年数据的总体精度和大豆、玉米的制图精度、用户精度和F1分数。

分析数据可以得到，单年份数据训练的4个模型在美国的3个测试区均能实现当年空间泛化，3个区域总体精度达到80%以上，玉米的制图精度为



83%—87%，F1 分数在 0.8—0.84，大豆的制图精度为 73%—85%，F1 分数在 0.78—0.84，略低于玉米；在黑河市测试区，4 个模型的总体精度较美国测试区精度降低，介于 60%—77%，玉米和大豆的制图精度也明显低于美国测试区域。

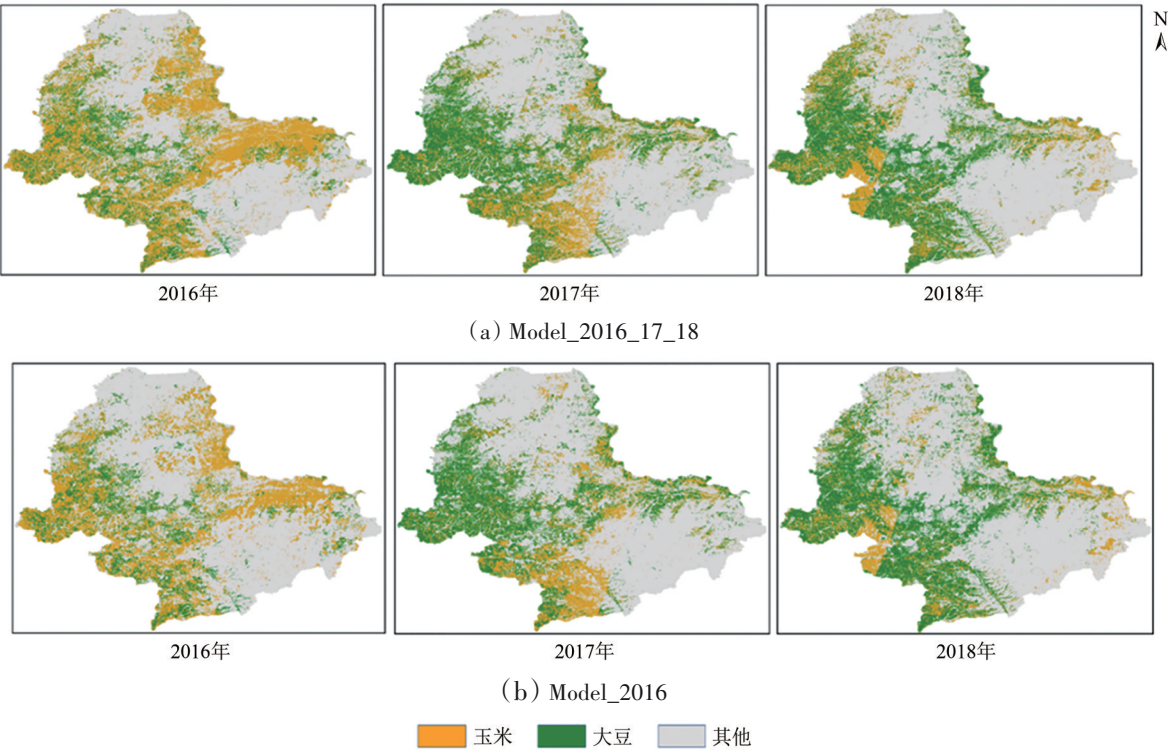


图6 Model\_2016\_17\_18 和 Model\_2016 在黑河市的玉米、大豆预测结果  
Fig. 6 Maize and soybean prediction results of Model\_2016\_17\_18 and Model\_2016 in Heihe City

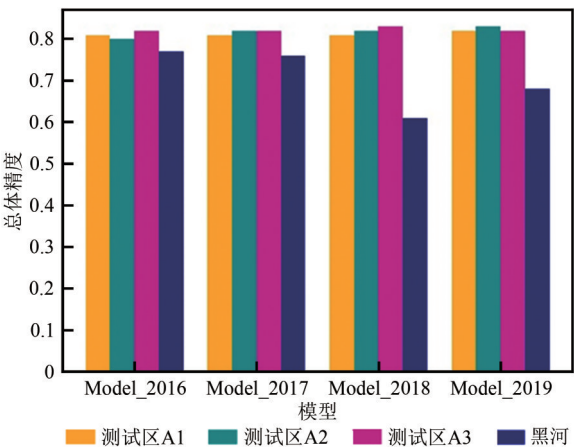


图7 不同模型在测试区预测当年数据的总体精度  
Fig. 7 Overall accuracy of different models in the test area in predicting the current year's data

4.3 分类模型时间泛化

4.3.1 美国测试区多年泛化

为了验证模型在时间尺度泛化能力，以美国

某一年的训练模型为基模型，预测测试区内其他年份的作物。图 8 是 6 个模型在美国 3 个测试区的作物预测结果。4 个单年份的作物识别模型在美国测试区的总体精度没有大的变化，介于 80%—83%，F1 分数在 0.80—0.87；大豆的 F1 分数 0.76—0.84。多年数据训练的作物识别模型 model\_2016\_17 和 model\_2016\_17\_18 在美国 3 个测试区的总体精度较单年份并没有明显提升，最高达到 84%，玉米 F1 分数最高达到 0.88，大豆 F1 分数最高 0.84，说明多年份数据训练的作物识别模型在美国测试区对于作物识别时间泛化精度提升并不明显。从不同年份的时间泛化角度分析，在美国测试区，这些模型并没有表现出时间泛化上的显著差异，说明在美国测试区，无论是单年份，还是多年份的作物分类模型，均能实现跨年份的作物模型泛化。

表 4 不同模型在测试区域预测当年玉米和大豆精度  
Table4 Accuracy of prediction of maize and soybean by different models in the test area in the current yea

测试区域	模型名称 (预测年份)	玉米			大豆			测试区域	模型名称 (预测年份)	玉米			大豆		
		制图精度	用户精度	F1分数	制图精度	用户精度	F1分数			制图精度	用户精度	F1分数	制图精度	用户精度	F1分数
A1	model_2016(2016)	0.85	0.81	0.83	0.80	0.79	0.79	A3	model_2016(2016)	0.87	0.81	0.84	0.79	0.83	0.81
	model_2017(2017)	0.84	0.80	0.82	0.80	0.79	0.79		model_2017(2017)	0.86	0.81	0.83	0.79	0.83	0.81
	model_2018(2018)	0.84	0.81	0.82	0.74	0.83	0.78		model_2018(2018)	0.85	0.78	0.81	0.75	0.85	0.80
	model_2019(2019)	0.83	0.81	0.82	0.85	0.84	0.84		model_2019(2019)	0.82	0.80	0.81	0.75	0.85	0.80
A2	model_2016(2016)	0.85	0.78	0.81	0.75	0.84	0.79	黑河市	model_2016(2016)	0.76	0.82	0.79	0.62	0.70	0.66
	model_2017(2017)	0.82	0.79	0.80	0.76	0.84	0.80		model_2017(2017)	0.72	0.69	0.70	0.63	0.65	0.64
	model_2018(2018)	0.85	0.84	0.84	0.76	0.82	0.79		model_2018(2018)	0.65	0.67	0.66	0.57	0.67	0.62
	model_2019(2019)	0.84	0.80	0.82	0.73	0.84	0.78		model_2019(2019)	0.63	0.62	0.62	0.61	0.67	0.64

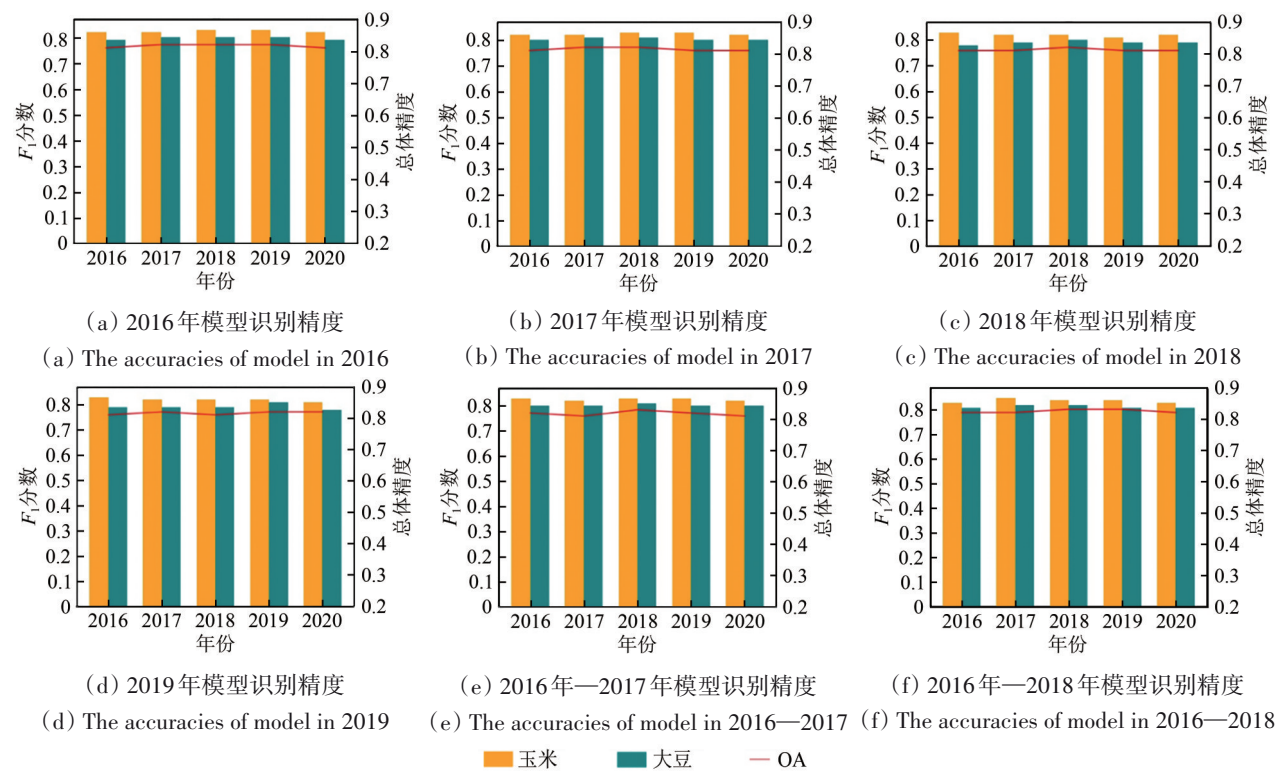


图 8 美国测试区玉米、大豆 F1 分数和总体精度

Fig. 8 F1 scores and overall accuracy of maize and soybean in US test area

4.3.2 黑河测试区多年泛化分析

图 9 是 6 个作物分类模型在黑河市进行玉米、大豆的识别结果。在 4 个单年份模型的结果中，每年的总体精度为 69%—79%，玉米 F<sub>1</sub> 分数在 0.69—0.80；大豆的 F<sub>1</sub> 分数低于玉米，在 0.65—0.73。两个多年份的作物识别模型总体精度得到了提升，为 72%—79%，玉米 F1 分数在 0.69—0.81；大豆 F1 分数在 0.66—0.76。单年份模型在黑河测试区当年数据上表现的更好，但时间泛化到其他年份表现略差。

由于黑河市 2018 年、2019 年云覆盖严重造成有效遥感影像数据缺失过多，导致这两个年的预测结果精度较低，因此对这两年结果不再做深入分析。

5 讨论

5.1 特征对模型迁移能力的影响分析

深度学习作为一种“端—端”的机器学习方法，特征不变性是模型泛化的基础。为深入探究深度学习模型的时空泛化能力，提取并分析 2016 年

训练区和测试区的影像特征, 图 10 训练数据和测试数据两个时期的特征, 波段组合为 Swir2、Swir1 和 Nir。在第一时期数据中, 作物刚播种, 玉米和大豆的光谱特征接近“裸地”, 二者无明显区别,

这一时期影像可以与处于生长的植被有效地区分开; 在第二时期数据中, 玉米和大豆的光谱差异明显, 这是进行二者分类的遥感特征基础。

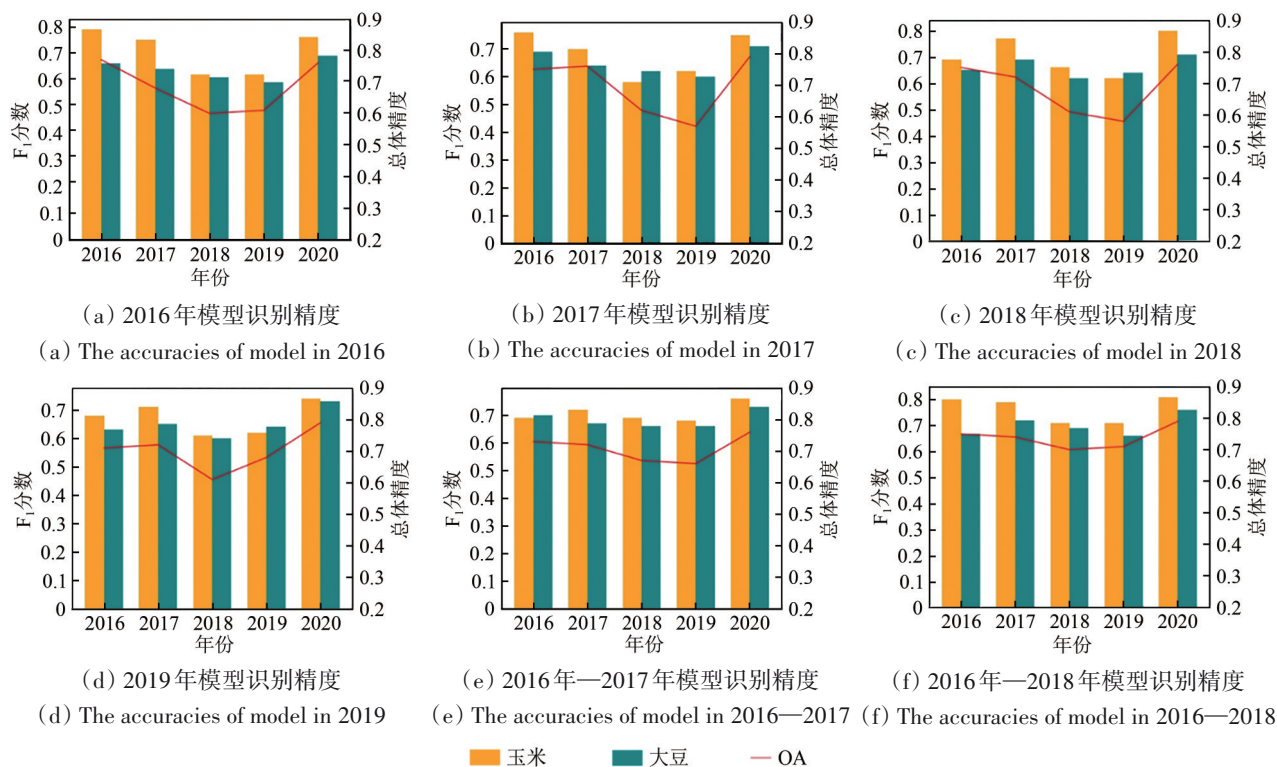


图9 黑河市测试区玉米、大豆F1分数和总体精度

Fig. 9 F1 fractions and overall accuracy of maize and soybean in the test area of Heihe City

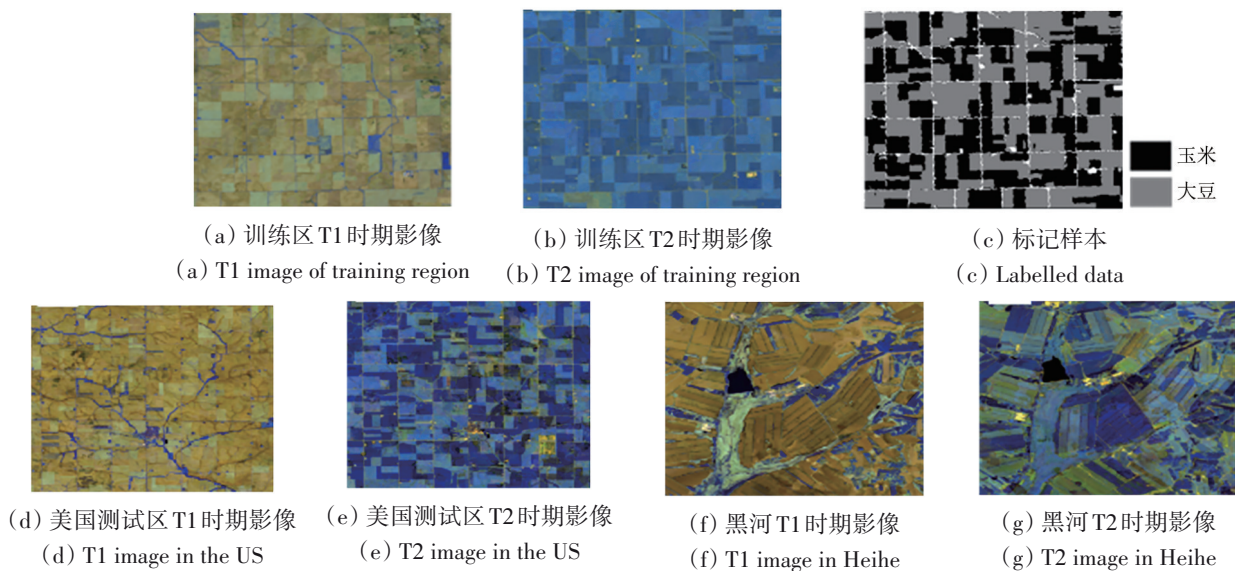


图 10 训练数据和测试数据的光谱特征

Fig. 10 Spectral characteristics of training data and test data



本研究为进一步研究数据光谱特征分布, 将2016年美国训练数据和测试区域A1的数据光谱特征进行统计分析, 见图11和12。训练数据的玉米和大豆的光谱值约在500—800具有可分性, 与CDL数据在测试区域对应的光谱值一致, model\_2016对A1的预测结果中玉米和大豆的光谱集中在这个范围内, 但相近光谱值内大豆和其他类别的像元数量与训练区的差别较大, 因此大豆和其他类别的精度较低。Muhammad等(2015b)使用历史参考数据训练出的模型跨年份进行泛化推广, 证实邻近年份泛化能力强, 从美国、黑河之间泛化结果可以看到模型泛化能力跟年份的相关是因为年际间作物遥感特征相似, 因此在年际间模型进行泛化应用时, 确保年际间的遥感特征具有一

定的一致性和稳定性。

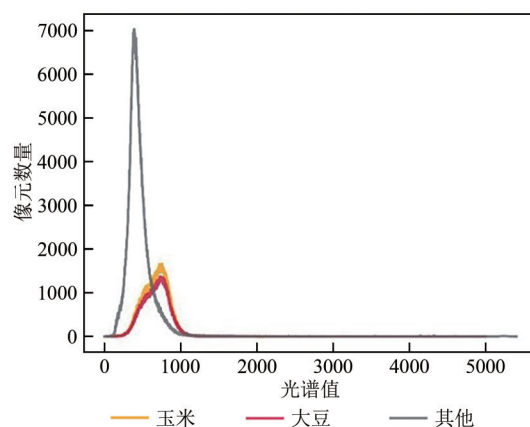
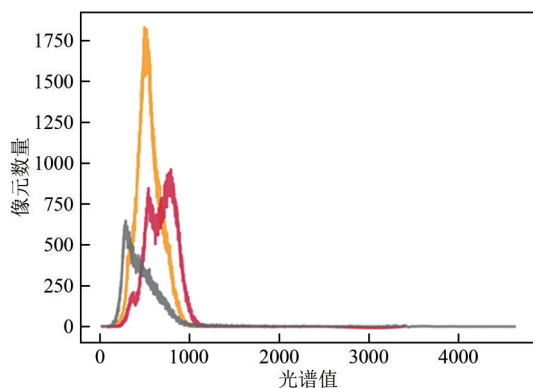


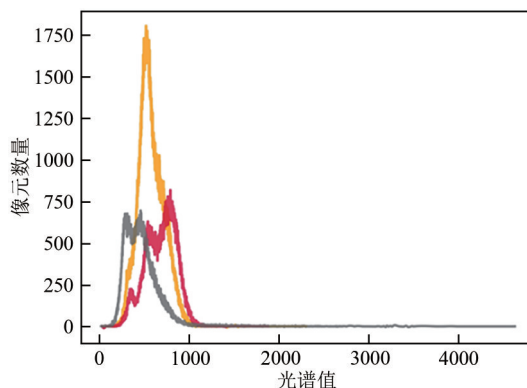
图11 美国2016年训练数据光谱特征分布图

Fig. 11 Spectral feature distribution of US training data in 2016



(a) 2016年CDL标签对应光谱特征

(a) The corresponding spectral feature of 2016 CDL data



(b) Model\_2016预测结果光谱特征

(b) The spectral feature of 2016 Model-based predicted result

— 玉米 — 大豆 — 其他

图12 2016年A1数据光谱特征

Fig. 12 Spectral characteristics of data in A1 region in 2016

深度学习工作机制复杂, 参数庞大, 常被称为“黑箱”算法, 能够对其工作机制和运算原理进行分析, 增强其可解释性, 可提高模型的泛化能力(Zintgraf等, 2017; 纪守领等, 2019), 分析模型的最后一层特征是检验模型泛化能力的一种常用作法。本研究可视化U-net解码过程的最后一层“conv\_out\_new”, 通过提取的特征解释模型的泛化能力。

图13是美国测试区的部分区域及其响应特征可视化结果, 图13(d)和图13(e)可见特征激活的高响应值区域对应CDL中的大豆和玉米, 在图13(d)中模型对于玉米的响应高, 在图13(e)中对应大豆的激活响应特征, 而在图13(f)中较

高激活响应值的地物更多为各类地物边界和道路, 最后图13(g)和图13(h)反映这些层对大豆和与玉米的激活响应低。

图14是黑河测试区部分区域及其响应特征可视化图, 图14(a)和图14(b)分别是测试区的两个时期数据, 图14(c) — (g)是提取分类模型上采样最后一层的特征可视化结果。分析其特征可视化结果发现, 在图14(d)中模型表现为高响应值的区域对应“玉米”, 图14(e)中高响应区对应“大豆”, 图14(f)对应其他地物类型, 但是对于其他地物的边界也表现为高响应值, 但存在误分, 同样图14(f)和图14(g)也是上述地物的低响应。

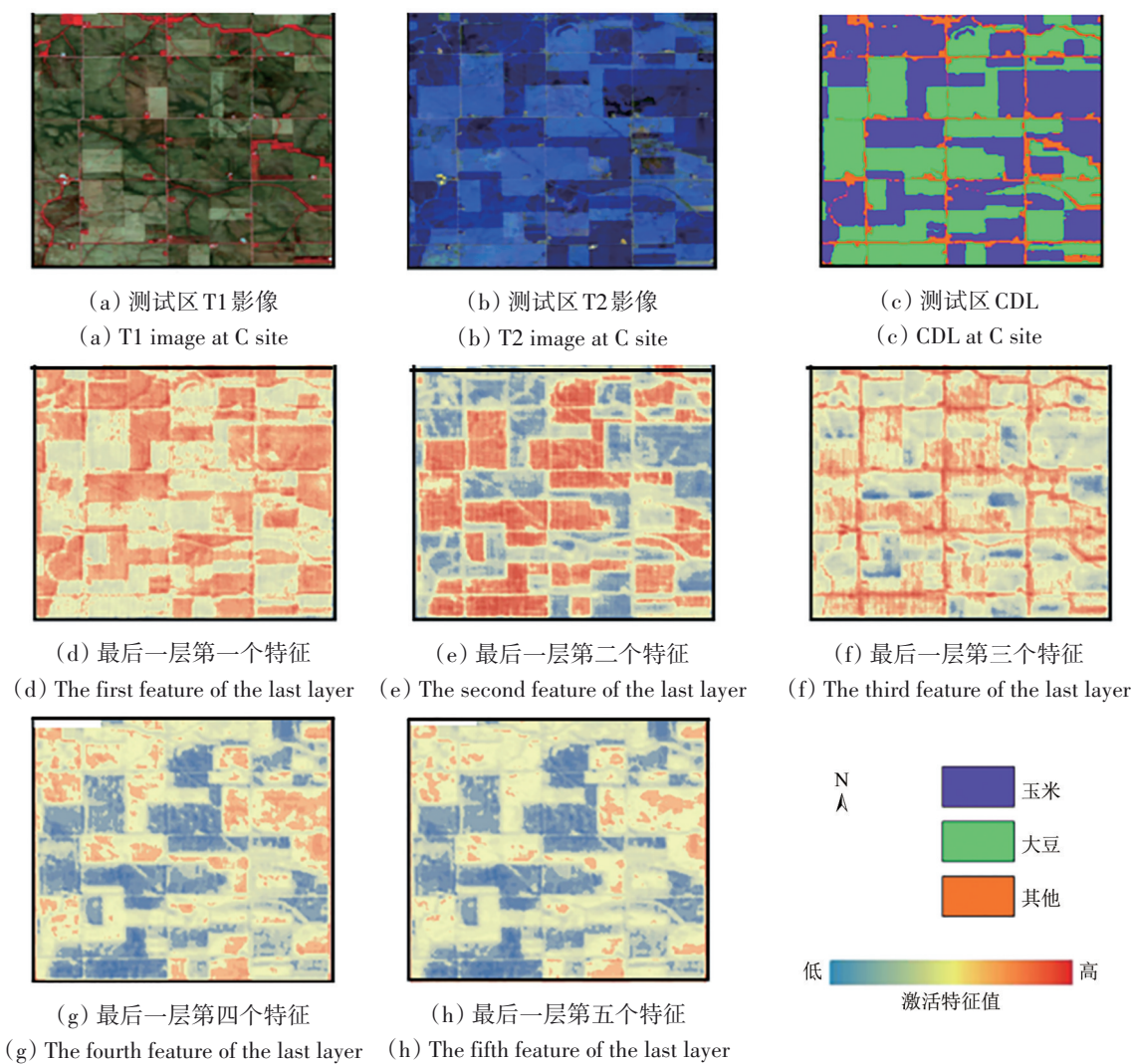
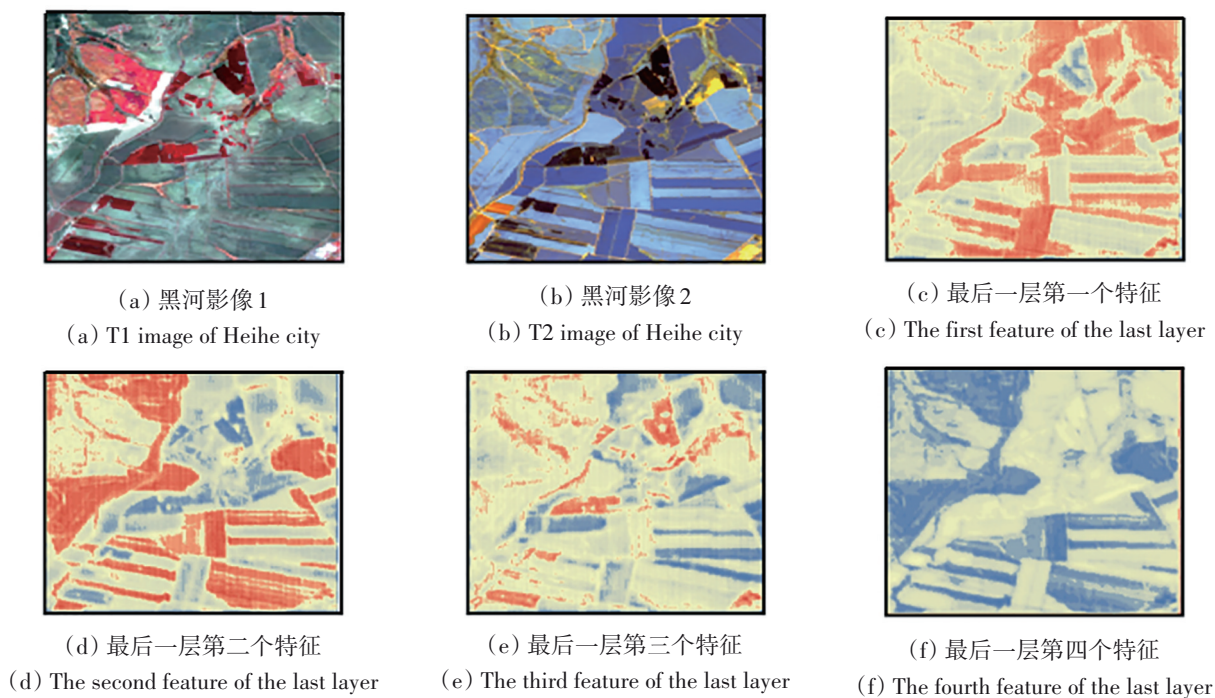


图 13 美国测试区特征可视化图

Fig. 13 Visualization diagram of extracted features in American test area



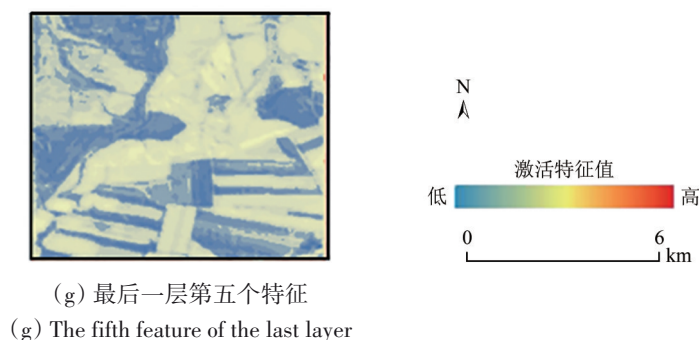


图14 黑河市测试区提取W可视化图

Fig. 14 Visualization diagram of extracted features in the test area of Heihe City

分析U-net模型架构，上采样的过程不仅还原了输入数据的分辨率和空间维度信息，同时结合上下文信息，但是在最后一层卷积后会丢失一些细小特征。结合精度评价和特征可视化结果分析，该模型在美国和中国两个测试区的作物特征描述是比较准确的，但是在中国测试区存在其他地物误分和边界不够清晰的问题，最大原因可能是美国和中国两个测试区的田块尺寸和作物种植结构差异。

## 5.2 与随机森林时空泛化对比分析

随机森林 (Random Forest RF) 是一种具有较高准确度、良好抗噪能力的传统机器学习算法，被许多领域广泛应用 (Smith 等, 2010)，其本质是基于 bagging 框架的多棵决策树的集成学习方法，它采用 bootstrap 抽样方法从原始训练集中抽取多个训练子集，每个训练子集都运用一定算法进行节点分裂，从而构建对应的一棵决策树，所有训练子集构建的决策树组合即为随机森林，预测时基于每个像元每棵树都输出一个结果进行投票，投票数最多的类别为对应像元的最终结果。

本研究使用2017年的训练数据训练随机森林，预测数据为2016年—2020年4个测试区域的数据，数据集的组织与深度学习保持一致。分类体系与深度学习模型一致，利用GEE中的 stratifiedSample 方法将CDL数据转化为点数据，利用 sampleRegions 构建训练数据集，将数据和训练样本输入到 scikit-learn 进行训练模型和结果预测，子决策树数量  $n=300$ ，其余参数默认，得到作物分类结果之后，进行精度评价。随机森林的预测结果如图15所示。

结合图7、图8和图15分析得到，随机森林的预测结果较深度学习模型低，美国预测结果总体

精度最高为75%，黑河市的预测结果总体精度最高为62%，说明随机森林的时空泛化能力较深度学习作物分类模型更差。随机森林是基于数据的统计学习模型，较深度学习，随机森林没利用除光谱之外的深层信息，如纹理、空间等特征，导致模型鲁棒性差，不合适时空泛化推广。

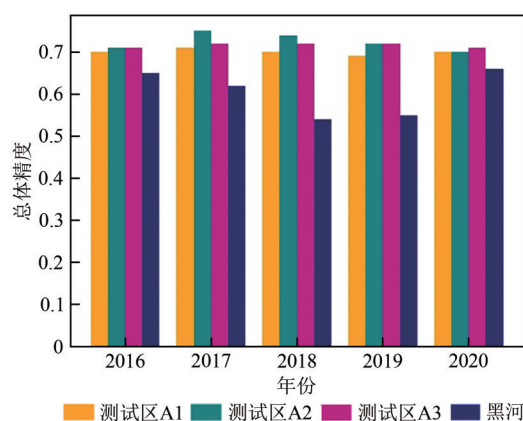


图15 随机森林时空泛化结果

Fig. 15 Spatial and temporal generalization results of random forest

## 5.3 测试集时间窗口选择对泛化的影响

为了进一步探讨测试集数据窗口大小对泛化效率的影响，选取黑河市嫩江县作为泛化区域，选取测试数据时缩小时间窗口至一个半月内，使用 model\_2016\_17\_18 进行作物识别。研究发现当缩小数据合成的时间窗口时，该模型的预测结果总体精度均提升到79%，且大豆的识别精度提升更为明显 (表5)，提升至75%，说明适当缩小影像的时间窗口有助于提升作物识别的精度，究其原因是预测数据的合成时间窗口缩小能更加精准的刻画作物生长的关键特征。Wang 等 (2019) 研究通过减少不同州和年份的标签数据来体现样本均



衡问题,发现在GDD (Growing Degree Days)相似的地区和年份上训练的随机森林模型迁移到目标区域的作物识别准确率始终超过80%,同时准确率会随着GDD差异的增大而降低,说明了提高训练区和测试区遥感特征的一致性是提高精度的关键,与本文发现相同。然而这方面的研究仍比较少,如何在大区域尺度通过物候匹配提高模型泛化能力尤为关键,将会是一个重点研究方向。

表5 model\_2016\_17\_18预测嫩江县2017年数据精度评价

Table 5 Model\_2016\_17\_18 predict data accuracy evaluation in Nenjiang County in 2017

窗口大小	OA	F1	
		玉米	大豆
缩小时间窗口前	0.75	0.72	0.69
缩小时间窗口后	0.79	0.77	0.75

## 6 结 论

本研究以美国伊利诺伊州和印第安州作为训练数据的采集区域,历史CDL数据作为作物标签,训练了2016年—2019年4个单年份数据的玉米、大豆分类模型,两个混合年份数据的分类模型,在美国和中国设置了4个空间泛化区域,探究了基于CDL迁移训练深度学习模型进行作物分类的时空泛化能力,并与随机森林分类模型进行了对比,得到以下结论:

(1) 本研究的6个秋粮分类模型在美国的空间泛化精度优于中国,美国3个测试区秋粮分类精度无空间差异。通过对模型提取到的特征和训练测试数据光谱特征进行分析之后发现玉米和大豆在第二时期数据中的光谱特征差异是其分类的基础,在美国和中国出现空间泛化差异的原因是两个区域之间作物种植结构和地块景观的差异,因此模型在泛化时对分类特征的提取出现了空间差异。

(2) 本研究的6个秋粮分类模型在美国3个测试区2016年—2020年的时间泛化总体精度在80%以上,玉米的精度高于大豆;在黑河市测试区的时间泛化能力总体低于美国,两个区域之间相差约8%,但多年份分类模型优于单年份分类模型,玉米的精度优于大豆。多年份模型能够提升模型时空泛化能力的原因是多年的样本扩大了标识样本数量,将中国与美国秋粮种植的差异通过样本数量的增加进行弥补。

(3) 随机森林秋粮分类模型的时空泛化的总

体精度较深度学习在美国测试区低7%,在黑河市测试区低8%。随机森林是基于数据的统计学习模型,与深度学习相比,随机森林未利用光谱之外的深层信息,且鲁棒性不强,因此时空泛化能力差。

本研究已经证实了利用CDL作为标签样本训练模型,可以实现空间、时间的泛化。在今后研究中将会对模型泛化数据的最适宜时间窗口和跨区域泛化时农业景观和作物种植结构的影响做进一步探究。同时现在作物分布的数据集越来越多,这些数据集能够进一步提供丰富的作物标签数据用于深度学习模型训练,为大尺度作物深度学习识别标签样本缺乏提供解决途径。

## 参考文献(References)

- 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]
- Egli D B. 2008. Comparison of corn and soybean yields in the United States: historical trends and future prospects. *Agronomy Journal*, 100(S3): S-79-S-88 [DOI: 10.2134/agronj2006.0286c]
- Hu Q, Wu W B, Song Q, Yu Y, Yang P and Tang H J. 2015. Recent progresses in research of crop patterns mapping by using remote sensing. *Scientia Agricultura Sinica*, 48(10): 1900-1914 (胡琼, 吴文斌, 宋茜, 余强毅, 杨鹏, 唐华俊. 2015. 农作物种植结构遥感提取研究进展. *中国农业科学*, 48(10): 1900-1914) [DOI: 10.3864/j.issn.0578-1752.2015.10.004]
- Ji S L, Li J F, Du T Y and Li B. 2019. Survey on techniques, applications and security of machine learning interpretability. *Journal of Computer Research and Development*, 56(10): 2071-2096 (纪守领, 李进锋, 杜天宇, 李博. 2019. 机器学习模型可解释性方法、应用与安全研究综述. *计算机研究与发展*, 56(10): 2071-2096) [DOI: 10.7544/j.issn1000-1239.2019.20190540]
- Jiang T, Liu X N and Wu L. 2018. Method for mapping rice fields in complex landscape areas based on pre-trained convolutional neural network from HJ-1 A/B data. *ISPRS International Journal of Geo-Information*, 7(11): 418 [DOI: 10.3390/ijgi7110418]
- Konduri V S, Kumar J, Hargrove W W, Hoffman F M and Ganguly A R. 2020. Mapping crops within the growing season across the United States. *Remote Sensing of Environment*, 251: 112048 [DOI: 10.1016/j.rse.2020.112048]
- Liu J, Wang L M, Yang F G, Yao B M and Yang L B. 2018. Recognition ability of red edge and short wave infrared spectrum on maize and soybean. *Chinese Agricultural Science Bulletin*, 34(35): 120-

- 129 (刘佳, 王利民, 杨福刚, 姚保民, 杨玲波. 2018. 红边与短波红外谱段的玉米大豆识别能力研究. 中国农学通报, 34(35): 120-129)
- Ma L, Xu X G, Jia J H, Huang W J, Liu L Y and Cheng Y P. 2008. Crop classification method using multi-temporal TM images. Transactions of the Chinese Society of Agricultural Engineering, 24(S2): 191-195 (马丽, 徐新刚, 贾建华, 黄文江, 刘良云, 程一沛. 2008. 利用多时相TM影像进行作物分类方法. 农业工程学报, 24(S2): 191-195) [DOI: 1002-6819(2008)-(Supp.2) -0191-05]
- Muhammad S, Niu Z, Wang L, Aablikim A, Hao P Y and Wang C Y. 2015b. Crop classification based on time series MODIS EVI and ground observation for three adjoining years in Xinjiang. Spectroscopy and Spectral Analysis, 35(5): 1345-1350 [DOI: 10.3964/j.issn.1000-0593(2015)05-1345-06]
- Pan Z K, Xu J S, Guo Y B, Hu Y M and Wang G X. 2020. Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. Remote Sensing, 12(10): 1574 [DOI: 10.3390/rs12101574]
- Phalke A R, Özdoğan M, Thenkabail P S, Erickson T, Gorelick N, Yadav K and Congalton R G. 2020. Mapping croplands of europe, middle east, russia, and central asia using landsat, random forest, and google earth engine. ISPRS Journal of Photogrammetry Remote Sensing, 167: 104-122 [DOI: 10.1016/j.isprsjprs.2020.06.022]
- Ronneberger O, Fischer P and Brox T. 2015. U-net: convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4\_28]
- Smith A, Sterba-Boatwright B and Mott J. 2010. Novel application of a statistical technique, Random Forests, in a bacterial source tracking study. Water Research, 44(14): 4067-4076 [DOI: 10.1016/j.watres.2010.05.019]
- Song D J. 2019. Method Research on Extracting Spatial Distribution Information of Winter Wheat using Convolutional Neural Network. Taian: Shandong Agricultural University (宋德娟. 2019. 基于卷积神经网络提取冬小麦空间分布信息的方法研究. 泰安: 山东农业大学)
- Sun Z H, Di L P, Fang H and Burgess A. 2020. Deep learning classification for crop types in north dakota. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13: 2200-2213 [DOI: 10.1109/JSTARS.2020.2990104]
- Wang S, Azzari G and Lobell D B. 2019. Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques. Remote Sensing of Environment, 222: 303-317 [DOI: 10.1016/j.rse.2018.12.026]
- Wei S S, Zhang H, Wang C, Wang Y Y and Xu L. 2019. Multi-temporal SAR data large-scale crop mapping based on U-Net model. Remote Sensing, 11(1): 68 [DOI: 10.3390/rs11010068]
- Xie D F, Zhang J S, Pan Y Z, Wun P J and Yuan Z M Q. 2015. Fusion of MODIS and Landsat 8 images to generate high spatial-temporal resolution data for mapping autumn crop distribution. National Remote Sensing Bulletin, 19(5): 791-805 (谢登峰, 张锦水, 潘耀忠, 孙佩军, 袁周米琪. 2015. Landsat 8和MODIS融合构建高时空分辨率数据识别秋粮作物. 遥感学报, 19(5): 791-805) [DOI: 10.11834/jrs.20154213]
- Xu J F, Zhu Y, Zhong R H, Lin Z X, Xu J L, Jiang H, Huang J F, Li H F and Lin T. 2020. DeepCropMapping: a multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. Remote Sensing of Environment, 247: 111946 [DOI: 10.1016/j.rse.2020.111946]
- Xu Q, Li Q L, De Vroey M, Zhang W P and Fan J L. 2019. Multi-crop synchronization identification based on random forest. Shandong Agricultural Sciences, 51(3): 135-139 (许淇, 李启亮, De Vroey M, 张吴平, 范锦龙. 2019. 基于随机森林算法的多作物同步识别. 山东农业科学, 51(3): 135-139) [DOI: 10.14083/j.issn.1001-4942.2019.03.028]
- Xu Q, Zhang J S, Zhang F, Gai S, Yang Z and Duan Y M. 2022. Applicability of weak samples to deep learning crop classification. National Remote Sensing Bulletin, 26(7): 1395-1409 (许晴, 张锦水, 张凤, 盖爽, 杨志, 段雅鸣. 2022. 深度学习农作物分类的弱样本适用性. 遥感学报, 26(7): 1395-1409) [DOI: 10.11834/jrs.20221127]
- Yang C H, Everitt J H and Murden D. 2011. Evaluating high resolution SPOT 5 satellite imagery for crop identification. Computers and Electronics in Agriculture, 75(2): 347-354 [DOI: 10.1016/j.compag.2010.12.012]
- You N S and Dong J W. 2020. Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine. ISPRS Journal of Photogrammetry and Remote Sensing, 161: 109-123 [DOI: 10.1016/j.isprsjprs.2020.01.001]
- Zhang X, Wu B F, Ponce-Campos G E, Zhang M, Chang S and Tian F Y. 2018. Mapping up-to-date paddy rice extent at 10 m resolution in China through the integration of optical and synthetic aperture radar images. Remote Sensing, 10(8): 1200 [DOI: 10.3390/rs10081200]
- Zhao Y S. 2003. Principles and Methods of Remote Sensing Application Analysis. Beijing: Science Press (赵英时. 2003. 遥感应用分析原理与方法. 北京: 科学出版社)
- 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]
- Zhong Y F, Hu X, Luo C, Wang X Y, Zhao J and Zhang L P. 2020. WHU-Hi: UAV-borne hyperspectral with high spatial resolution ( $H^3$ ) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. Remote Sensing of Environment, 250: 112012 [DOI: 10.1016/j.rse.2020.112012]
- Zintgraf L M, Cohen T S, Adel T and Welling M. 2017. Visualizing deep neural network decisions: prediction difference analysis. [DOI:10.48550/arXiv.1702.04595].

## Transferring deep convolutional neural network models for generalization mapping of autumn crops

ZHANG Feng<sup>1,2,3</sup>, ZHANG Jinshui<sup>1,2,3,4</sup>, DUAN Yaming<sup>1,2,3</sup>, YANG Zhi<sup>1,2,3</sup>

1. State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China;
2. Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;
3. Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;
4. Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China

**Abstract:** Deep Convolutional Neural Networks (DCNNs) have been increasingly applied in remote sensing crop recognition due to their “end-to-end” advantages and efficient extraction of shallow shape details and deep semantic features. However, deep learning models require a large number of labeled samples, which are time-consuming, labor-intensive, and costly to obtain, limiting the 2016–2020 period. The U-net model based on CDL training can be popularized and applied in the United States. First, the overall accuracy of time generalization in the three test areas in the United States from 2016 to 2020 is more than 80%, and the recognition accuracy of corn is higher than that of soybeans. Deep learning models have good transferability in space. Second, for autumn grain in Heihe City, the average recognition accuracy of corn for many years is 3% higher than that of soybean. This is because the corn planting plot is more regular and the planting scale is higher than that of soybean; the overall accuracy of autumn grain identification in a single year is between 69% and 79%. The year classification model is better than the single-year classification model, which may be because the representativeness of the training samples is enhanced with the increase of the number of labeled samples, and the difference in autumn grain planting between China and the United States can be compensated by the expansion of the number of samples. However, the model is migrated to the Heihe region of China. The accuracy of the models is lower than that of the continental United States, which is due to the inconsistency of remote sensing response characteristics due to differences in intercontinental climate and crop planting habits. These, in turn, reduce the generalization performance of the model. The DCNN model is better than random forest algorithm because of the training process driven by big data. The principle of transferring the basic trained crop classification model to map crop distribution timely and accurately has broad prospects for application across a large scale of time and space. The consistency of remote sensing features and phenology of the crops of the test area compared to the training data are fundamental factors that must be carefully considered, as these determine the success of high crop mapping performance. Therefore, it is essential to analyze the prerequisites when transferring the model to other places.

**Key words:** remote sensing, transfer learning, CDL, time-space generalization, soybeans, maize

**Supported by** National Natural Science Foundation of China Major Program (No. 42192580, 42192584); Major Project of High Resolution Earth Observation System (Civil Part) (No. 20-Y30F10-9001-20/22)