

MPSPNet 和 UNet 网络下山东省高分辨耕地遥感提取

李倩楠^{1,3} 张杜娟^{1,3} 潘耀忠^{*1,2,3} 代佳佳^{1,3}

1. 遥感科学国家重点实验室, 北京师范大学地理科学学部, 北京 100875;

2. 青海师范大学地理科学学院, 西宁 810016;

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

摘要: 高分辨率遥感影像中耕地特征复杂, 人工目视解译和传统的遥感影像分类方法提取能力有限, 无法实现大范围的自动化高精度耕地提取。深度学习技术因具有较强的地物表达能力, 在遥感影像信息自动提取方面表现出了优越的性能, 为大范围耕地的精细化自动提取提供了新的思路。探究不同典型网络模型在不同景观特征耕地提取上的适用情况对耕地提取质量和效率的提升具有重要意义。基于此, 本研究以高分一号及高分二号融合的 2 米分辨率数据为数据源, 采用改进的金字塔场景解析网络(Modified Pyramid Scene Parsing Network, MPSPNet)和 UNet 网络模型, 应用于山东省的耕地精细自动化提取, 并传统面向对象的方法对比, 探究两种深度卷积神经网络模型在大尺度耕地自动提取中的适用性。研究获得以下结论:(1) MPSPNet 模型和 UNet 模型在区/县尺度的耕地提取上性能优于传统的面向对象的分类方法, 在全省尺度的耕地提取上总体精度优于 90%且无明显差异。(2) 耕地景观特征是影响两模型耕地提取效果的重要因素, 模型的选择对耕地提取效果无明显影响。在耕地景观指数较低的地块规则平整的区域, 模型提取效果较好, 在耕地景观指数较高的地块破碎丘陵区域以及与耕地特征相近的地块区域, 模型提取效果较差, 并且 UNet 模型在这些区域误分耕地的概率更大。(3) 两模型在不同区域、不同时相的影像中能得到较好的耕地提取效果, 具有较强的泛化能力和时空迁移能力。

关键词: 耕地, 遥感, 卷积神经网络, MPSPNet, UNet

中图分类号: **文献标志码:**

1 引言

高精度大尺度的精细化耕地空间分布信息对耕地可持续发展、粮食安全等具有重要意义(Lu 等, 2017)。遥感影像技术的快速发展使得大量高分辨率遥感影像可为精细化的耕地等地物信息的准确提取提供良好的数据支撑(胡琮等, 2015), 但其自身的高度细节化、数据量大、类内差异大等因素给影像的自动解译带来了巨大的挑战, 在实际业务应用中以人工目视解译为主, 虽然这种方式得到的数据较为精确, 但不适用于大尺度范围的耕地提取。利用遥感影像分类方法提取耕地信息是一种更高效、全面和准确的方式。

传统遥感影像分类方法根据有无标签数据, 可以被分为非监督分类和监督分类。非监督分类方法主要通过聚类来实现, 即对各种波段的遥感影像的统计差异进行分析, 将具有不同统计属性的类分为不同的类

别。常用于耕地提取的非监督遥感影像分类方法有 K 均值聚类分析法和迭代自组织数据分析法 (ISODATA) (张明和黄双燕, 2019)。非监督分类方法受人为因素影响较小, 自动化程度较高, 但是由于遥感影像中耕地的特征复杂, 大量研究表明基于非监督分类的耕地信息提取效果不如监督分类(李鹏等, 2017)。监督分类方法是根据标签数据样本等先验知识对函数进行训练, 利用已经训练好的函数将未分类的对象与训练样本进行对比, 若相符, 则将其分类成耕地(Zhi 等, 2017)。常用于耕地提取的监督分类方法有最大似然法(Julien 等, 2011), 随机森林法(Hernandez 等, 2018), 支持向量机方法(李昌俊等, 2018), 决策树分类法(De 等, 2012), 神经网络算法等(张峰等, 2019), 虽然这些方法在耕地提取上的精度较高, 但是对影像的选取和分类中所利用的特征参数都有较高的要求, 无法获取影像中普适性高、精准可靠的分类特征(周培诚等, 2021)。此外, 由于高分辨率影像数据光谱信息特征

收稿日期: 预印本:

基金项目: 1. “十三五”国家重点研发专项: 地震风险暴露模型研究(编号: 2018YFC1504603)

2. 国家高分辨率对地观测系统(民用部分) 重大专项(编号:E03071112)

第一作者简介: 李倩楠, 1996 年生, 女, 硕士研究生, 研究方向为农业遥感。E-mail: lqn@mail.bnu.edu.cn

通信作者简介: 潘耀忠, 1965 年生, 男, 教授, 研究方向为农业遥感。E-mail: pyz@bnu.edu.cn

较少, 光谱信息的重叠和混杂现象严重, 因而基于像素光谱特征的分类方法无法满足大范围的精细化耕地提取。面向对象分类方法进行遥感影像信息提取时, 以多个邻近像元构成的对象为基本单元, 综合高分辨率遥感影像的光谱、形状、纹理特征等进行分类。众多研究表明, 面向对象的分类比传统的基于像素的分类更适于高分辨率影像信息的提取(Goodin 等, 2015; 王卫红等, 2011)。面向对象分析技术的耕地提取(Blaschke, 2010; Lu 等, 2007)虽能利用高分影像中耕地的光谱、形状和纹理等信息, 解决传统基于像元信息提取中精度不高、椒盐现象的问题, 但是该方法只是利用影像的浅层特征, 而忽略了分割对象之间的高层次特征, 往往导致错分、漏分, 鲁棒性差。同时这些浅层特征只是基于特定时间、特定位置设计, 扩展性差(Cheng 等, 2017), 难以在大尺度上推广应用。综上所述, 以上传统的耕地提取方法特征提取能力有限, 无法实现大尺度上高分辨率遥感影像的自动解译和耕地信息提取。

近年来, 深度学习方法通过训练大量样本数据, 自动学习提取主要特征, 实现自动化信息提取, 无须人工涉及且对于复杂多变的情况具有更好的鲁棒性, 在计算机视觉等领域取得了很好的效果, 是一种非常有价值的人工智能算法(Ma 等, 2019)。卷积神经网络(Cvolutional Neural Network, CNN)是深度学习中最著名、应用最广泛的算法(Lecun 等, 1989), 具有局部连接、权值共享等特性, 可以降低网络的复杂度, 减少参数数量, 性能远优于传统机器学习算法(Larochelle 等, 2009), 目前已被广泛应用于图像处理领域。LeNet-5 (LeCun 等, 1998)、AlexNet (Krizhevsky 等, 2012)、GoogleNet (Szegedy 等, 2015)、VGG (Simonyan 等, 2014)及 ResNet (He 等, 2016)是经典的 CNN 模型, 这些模型设计了深层级的卷积层对训练样本数据进行学习和特征信息提取, 模型的训练速度更快, 鲁棒性更强, 在图像分类方面有较好的表现。但是这些模型是通过将整幅影像切割成包含单一地物的若干子影像后, 分别对子影像中的地物进行识别, 无法实现像素级的分类, 并且存在高分辨率细节信息丢失严重的问题, 无法适用于高分辨率耕地的信息提取。在此基础上, Long 等人 (2015)提出了全卷积神经网络(Fully Convolutional Network, FCN), 将 CNN 中的全连接层换成了卷积操作, 从抽象的特征中恢复出每个像素所属的类别, 实现了遥感影像分类从图像级到像素级的扩展, 保留了图像的高分辨空间信

息, 减少了模型参数量, 提高了算法效率。

由于耕地类型复杂、作物种类多样、作物生长阶段不同等导致高分辨遥感影像中耕地的光谱特征多变、纹理特征复杂, 同物异谱现象严重, 特征提取难度较大, 而特征提取是遥感影像分类的基础, 对高分影像中耕地的特征挖掘是实现大尺度精细化耕地提取的关键。深度学习技术可以通过对高分影像中耕地的边界、纹理、上下文等丰富的多层次空间特征的学习与分析, 更加准确地挖掘高分辨率遥感影像中更深刻更本质的耕地特征, 提高耕地识别的准确度, 解决传统方法中泛化性差、鲁棒性差的问题(王桥等, 2021), 基于像素级分类的全卷积神经网络模型为大尺度、高分辨率耕地的自动化精确提取提供了新的思路。徐文娜 (2020)基于高分辨率遥感影像的研究证明了 UNet 网络在耕地提取上的学习效果和优越性。Zhang 等(2020)提出了一种改进的金字塔场景解析网络(Modified Pyramid Scene Parsing Network, MPSPNet)深度学习网络模型, 实现了在大尺度上进行耕地的高精度自动化提取。

虽然深度学习技术在遥感影像耕地提取上的可行性已经被证明, 但是对于典型算法在大范围耕地提取的能力差异, 不同的算法条件下, 模型对不同景观特征耕地提取的适用性以及深度学习模型在耕地提取上的泛化能力和迁移能力仍有待探究。基于此, 本研究以高分一号及高分二号合成 2 米分辨率遥感影像数据为数据源, 分别基于 MPSPNet 及 UNet 网络模型对影像中不同层次的耕地的空间特征、纹理特征等进行学习, 对山东省的耕地进行大尺度自动化提取, 探究典型深度学习算法在大范围耕地提取中的潜力。

2 实验区域及数据

2.1 实验区域概况

本研究以位于华北平原农业主产区的山东省为研究区, 共 137 个区/县, 总面积约 15.71 万平方千米, 境内耕地资源丰富, 类型多样, 耕地主要类型为旱地, 规整耕地主要分布在西南、西北及东部地区的平原及缓丘地区, 破碎耕地主要分布在中部丘陵地区。其丰富的耕地形态可以代表中国大部分耕地区域, 满足大尺度耕地自动提取适用性的研究需求。根据山东省地势情况及农业区划, 我们以区/县为单位, 将山东省划分为平原区和混合区(丘陵及平原混合区域)。

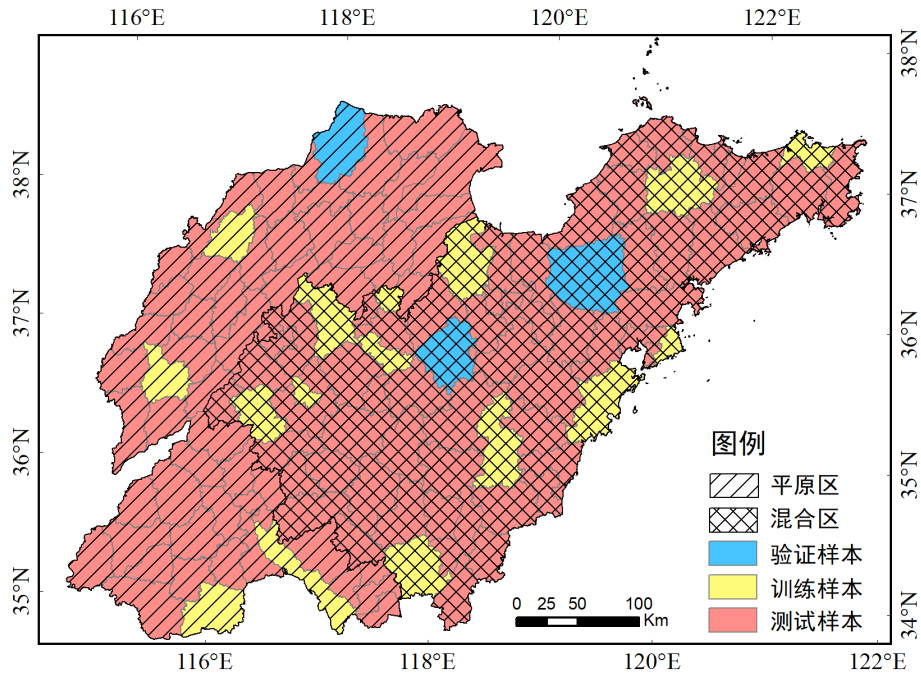


图 1 研究区及样本集分布

Fig.1 Study area and the distribution of samples

2.2 遥感影像数据

本研究使用的影像数据来源于高分应用综合信息服务平台(<http://gaofenplatform.com/index.html>), 高分卫星影像具备高空间分辨率、多光谱与高时间分辨率相结合的特点, 精度高、稳定度好, 具体载荷参数见表 1。本研究以区/县为单位, 镶嵌合成了覆盖山东省全部区/县的影像, 使用的影像以 2016 年 GF-1 号影像为主, 在 2016 年 GF-1 号影像不能覆盖的区域, 以 2016 年 GF-2 号影像、2015 年 GF-1 号影像及 2017 年 GF-1 号影像填补, 最终共使用了 346 景影像。在镶嵌时仅使用每景影像中无云/少云的部分, 影像投影为阿尔

伯圆锥形等面积投影, 所有影像时相分布如图 3 所示。在使用前为了使影像符合实验要求, 对其进行了配准、辐射校正、几何校正等预处理, 最后采用 Gram-Schmidt 光谱锐化融合方法 (Gram-Schmidt Pan Sharpening, GS) (Clayton 等, 1971; Laben.等, 2000)把全色波段和多光谱波段融合成 2m 分辨率的红绿蓝三波段影像, 目视检查判断融合后的影像地物清晰, 与原始多光谱图像无明显差异。此外, 考虑到存储空间和计算能力的限制, 本研究以区/县级行政单元为对象对影像进行裁剪、镶嵌和存储。全部影像数据基本覆盖耕地各个状态, 包含裸露的耕地以及农作物各个生长时期的耕地 (图 2)。

表 1 GF-1 号及 GF-2 号传感器参数信息

Table 1 Parameter information of GF-1 and GF-2 sensors

卫星	有效载荷	波段号	光谱范围 /um	空间分辨率 /m	幅宽/km	标准景尺寸 /km	侧摆能力 /(°)	全球覆盖 能力/天	重访时间 /天
GF-1	2m 分辨率全色 /8m 多光谱相机	B01	0.45-0.90	2	35(单台) 70(双台)	35×35	±35	41	4
		B02	0.45-0.52	8					
		B03	0.52-0.59						
		B04	0.63-0.69						
		B05	0.77-0.89						
GF-2	1m 分辨率全色 /4 米分辨率多光谱	B01	0.45-0.90	1	45(双台)	45×45	±35	69	5
		B02	0.45-0.52	4					
		B03	0.52-0.59						
		B04	0.63-0.69						
		B05	0.77-0.89						

2.3 标记样本数据集

本研究的标记样本数据集采用的地块图斑数据来源于第三次全国农业普查二类测量成果（中华人民共和国国家统计局和国务院第三次全国农业普查领导小组办公室，2016），为基于 2016 年 2 米分辨率的合成高分一号影像勾绘解译得到的矢量图层，总体精度高

于 98%，包括设施农业用地、耕地、园艺果树、林木、草地、水域、建筑物、道路、其他地物九大地物，各类型地块图斑如图 2 所示，本研究的耕地专指其中的农作物用地。首先，我们将原始矢量图层转为 2 米分辨率的栅格数据以和遥感影像相匹配，农作物用地标记为 1，非农作物用地标记为 2，最后与遥感影像数据三个波段叠加，作为最终的标记样本数据集。

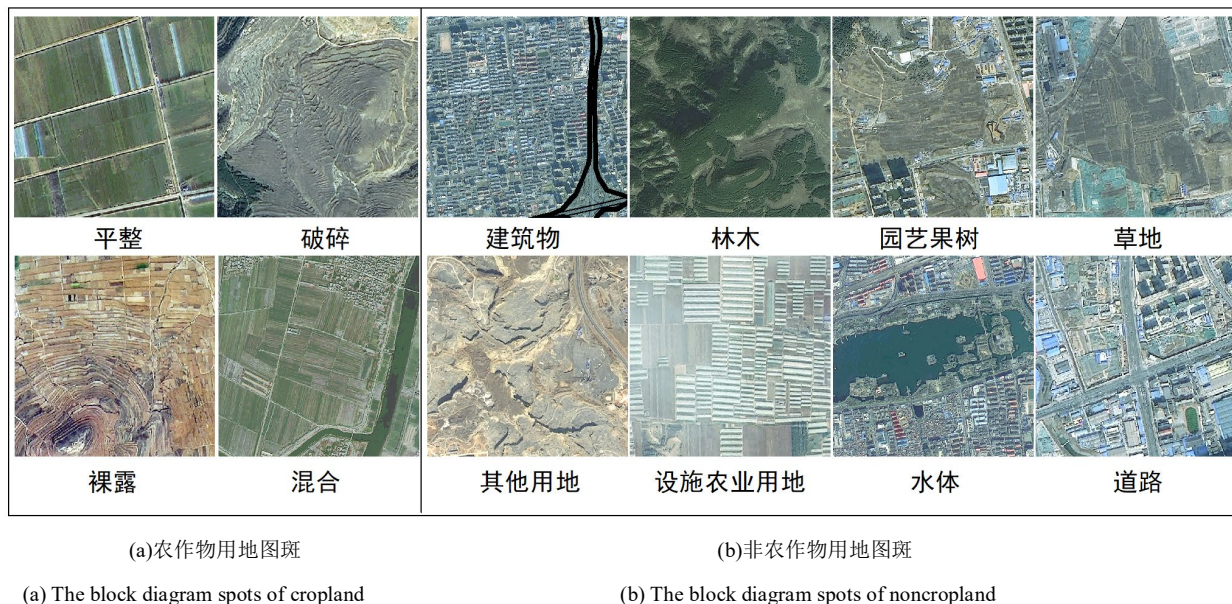


图 2 第三次全国农业普查地类图斑

Fig.2 The block diagram spots in the third national agricultural census

2.4 数据集划分

整个研究区域的标记样本数据集被分成三个数据集：训练样本数据集、验证样本数据集、测试样本数据集，其中训练样本用于训练模型，验证样本用于选择模型最优参数，测试样本用于最终评估模型性能。

在深度学习中，要训练得到一个高准确性的分类模型需要满足两个基本假设条件：(1)用于学习的训练样本与验证样本要满足独立同分布的条件；(2)要有足够多的有标记训练样本才能驱动学习得到一个准确的分类模型。因而标记样本的选择要同时满足以上两个条件并且要充分考虑待提取地物的空间分布等特征。基于以上原则，本研究以区/县为基本单元，参考 Ce 等（2018）和 Hugo 等（2017）的样本比例设置，同时考虑到本研究区的大面积覆盖和实际应用场景，只选取约 15%的区域来构建样本集进行模型的训练和选取，另外 85%的区域作为测试数据集以评估模型的性能。

在 15%区域中，有 80%的区域构建训练样本进行模型的训练，剩下 20%构建验证样本以进行最优模型参数的选取。综上最终选择 15 个区/县为训练样本集，3 个区/县为验证样本集，其余作为测试样本集。空间分布如图 1 所示，样本数量情况如表 2 所示，时间分布如图 3 所示。

本次研究所采用的深度学习框架为卷积神经网络框架 (Convolutional Architecture For Fast Feature Embedding, Caffe) (Jia 等, 2014)。为了避免样本产生边界效应及割裂，训练开始前分别对训练样本和验证样本采用 50%的重叠率进行切割，每块样本的大小为：256 像元×256 像元×4 波段，结果共产生 162013 块训练样本和 11877 块验证样本。将所有的训练样本数据和验证样本数据分别存储在轻量级内存映射数据库 (Lightning Memory-Mapped Database, LMDB) 中，影像数据和标签分别存储。

表 2 标记样本数据集数量

Table 2 The number of sample datasets with label

	区/县个数	面积(km ²)	面积占比(%)	样本块数
训练样本	15	21700.99	13	162013
验证样本	3	5648.79	3	11877
测试样本	119	140750.22	84	—

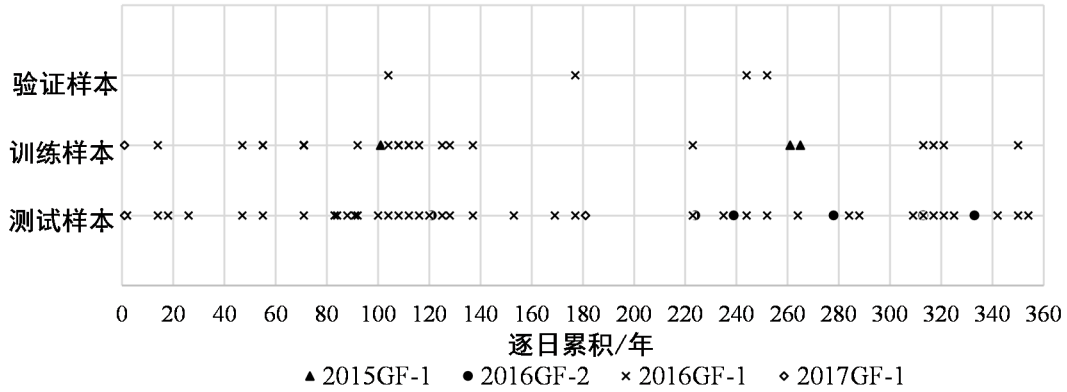


图 3 遥感影像/标记样本数据集时间分布（横坐标为一年中的累积日数，20 即为一年中的第 20 天）

Fig.3 The time distribution of the sample datasets with label /remote sensing images (Abscissa is the cumulative number of days in a year, 20 is the 20th day of the year)

3 实验方法与设计

3.1 基于深度学习模型的耕地提取方法

CNN 主要由输入层、卷积层、池化层、全连接层四部分组成。输入层为一幅图像的像素矩阵，卷积层是 CNN 中最为重要的部分之一，卷积层中每一个节点只和上一层中的部分神经元相连，通过不同的过滤器从神经元中提取更高层次的特征。池化层在卷积层之后，主要功能是进行特征选择，逐步减少特征空间的大小，减少参数个数和计算量，以减小过拟合的影响。全连接层在整个模型的最后部分，对池化层输出的特征进行分类操作。经过卷积层的信息提取以及全连接层的信息融合和分类计算，输出的结果是图像对应于某一类别的概率。

全卷积神经网络在卷积神经网络的基础上实现了端到端的像素级分类，并且其网络中的跳跃连接结构使模型具有较强的鲁棒性和精确性，可以用于准确学习和提取高分辨遥感影像中复杂的耕地特征和全局上下文信息，实现像素级的分类。UNet(Ronneberger 等, 2015)模型，金字塔场景解析网络 (Pyramid Scene Parsing Network, PSPNet) (Zhao 等, 2017)模型，DeepLabv3+ (Chen 等, 2018)模型等都是近年来流行的基于全卷积神经网络的像素级分类模型。

UNet(Ronneberger 等, 2015)网络是基于 FCN 的一个语义分割网络，它基于特征维度拼接实现特征融合，其中的跳跃连接结构实现了深层特征和浅层特征在不同尺度下的融合，加权交叉熵损失函数使模型具有分离边界的能力，并且模型复杂度低，结构简单灵活。PSPNet 解决了基于 FCN 模型中没有利用全局场景信息的问题，其中的深度残差网络结构基于卷积层可以实现影像特征的挖掘，网络中的金字塔池化模块基于提取的影像深层特征和浅层特征在多尺度上进行整合可以获取影像更丰富的全局信息。DeepLabv3+模型与 UNet 模型一样都是采用编码-解码器的结构，它基于空洞卷积融合多尺度特征信息，模型结构更加复杂，计算量更大。

高分辨遥感影像数据量大，特征复杂，并且大尺度影像中的耕地时相复杂、特征复杂，需要模型根据训练样本集学习影像中复杂的耕地边界等浅层特征以及耕地纹理、结构、上下文信息等深层特征，挖掘影像中耕地与非耕地的本质特征以快速、准确识别影像中的耕地。因而本研究采用模型结构相对简单、计算量小、能准确挖掘提取影像中深层特征和浅层特征，并能在不同尺度和区域下整合特征信息的 PSPNet 模型和 UNet 模型进行大范围的耕地提取。

3.1.1 MPSPNet 模型

MPSPNet 模型为在 PSPnet 模型基础上进行改进得到的。PSPNet 模型结构为深度残差网络架构 (Residual Networks, ResNet) 和金字塔池化模块 (Pyramid Pooling Module) 的结合。它可以解决用 CNN 进行图像解析、语义分割时没有利用整体场景信息而出现上下文匹配错误的问题, 可以融合合适的全局特征, 使最终的预测更加可靠。目前该方法在各种数据集上的应用已表现出较好的性能(Zhao 等, 2017), 对高级别特征的提取效果较为理想, 可应用于高分辨率影像耕地提取的研究。

ResNet 由 GoogLeNet 发展而来, 它创新性地设计了残差结构块(Residual building block), 可以降低参数的数目, 减少计算量, 同时解决其他网络中“随着网络加深, 准确率下降”的问题。金字塔池化模块(Pyramid Pooling Module)已经被证明是一个有效的能够整合不同层次区域信息从而获得全局信息的结构(Zhao 等, 2017)。

PSPNet 网络模型能完整地提取出高层次特征, 但随着卷积深度加深, 影像边缘信息会被模糊, 地物的边界信息无法被准确提取。因而本研究采用了 Zhang 等 (2020) 提出的一种改进的金字塔场景解析网络 (MPSPNet)深度学习网络模型, 该网络模型在 PSPNet 网络模型的基础上进行了改进, 即在深度残差网络 (ResNet)前加入一层卷积层用于浅层边界特征提取, 使模型能更精确地对耕地及其边界信息进行提取。MPSPNet 的模型网络架构如图 4 所示。图中 A 部分为卷积层浅层特征提取, 获取影像中耕地的边界特征, B 部分的 ResNet 结构对影像中耕地纹理、上下文等深度特征进行提取, 这里采用 8 个 ResNet 残差结构。C 部分为金字塔池化模块, 它通过对提取出的深层特征和浅层特征在不同的尺度上进行融合, 获取不同区域的上下文信息, 即影像中不同尺度的耕地空间特征信息。ResNet 输出的特征图层经过 4 层池化、卷积和上采样, 恢复至输入特征图像大小, 经过连接层与 ResNet 的深层特征图层连接。最后将带有不同子区域语义信息的深度特征和含有局部边界信息的浅层特征融合得到预测结果。

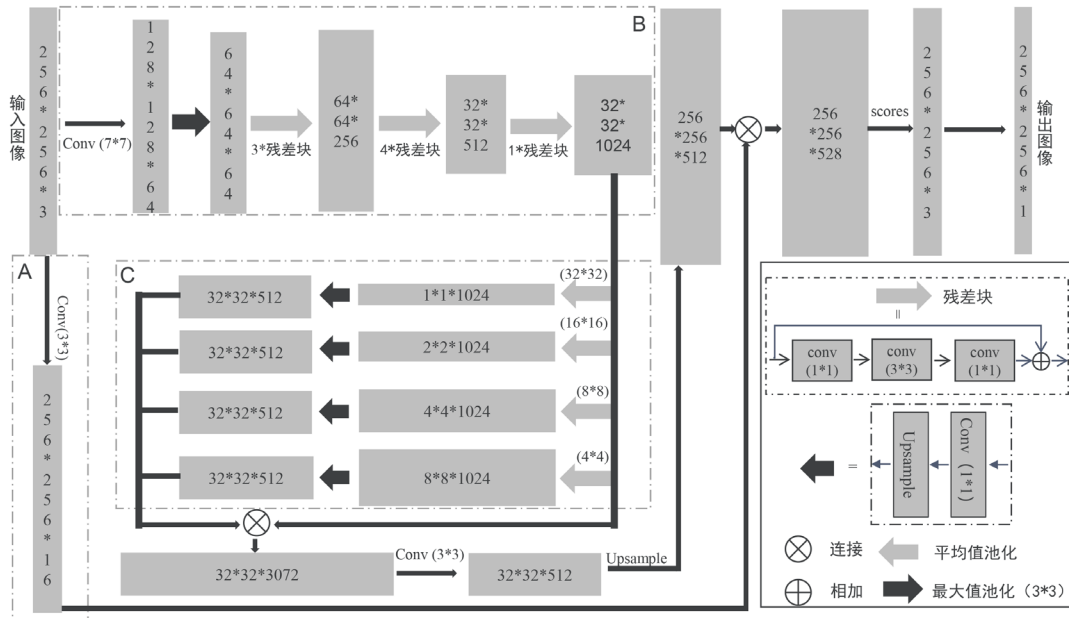


图 4 MPSPNet 网络架构图

Fig.4 The network architecture of MPSPNet

3.1.2 UNet 模型

UNet 模型为 U 型对称结构, 一侧是深层特征及浅层特征提取的卷积层, 另一侧是连接层及下采样层。UNet 网络的每个卷积层得到的特征图层都会连接到对应的下采样层, 从而实现每层特征图都有效使用到

后续计算中。同其他的一些网络结构不同, UNet 避免了直接在高级特征图层中进行监督和 loss 计算, 结合了低级特征图中的特征, 使得最终所得到的特征图结果中包含足够多的高级特征和低级特征 (Ronneberger 等, 2015), 实现不同尺度下特征的融合, 提高模型的分类结果精确度。

Unet 网络架构如图 5 所示。图中 A 部分为基于 ResNet 的影像纹理、上下文等深层特征的提取，使用了和 MPSPNet 模型中相同的 8 个 ResNet 残差结构块。图中 B 部分为连接层及下采样层，A 部分得到的耕地

深层和浅层特征图层通过连接层及下采样层全部连接，使影像中所有的耕地纹理等深层特征和边界等浅层特征被全部利用，最后经过卷积层生成最终的预测结果。

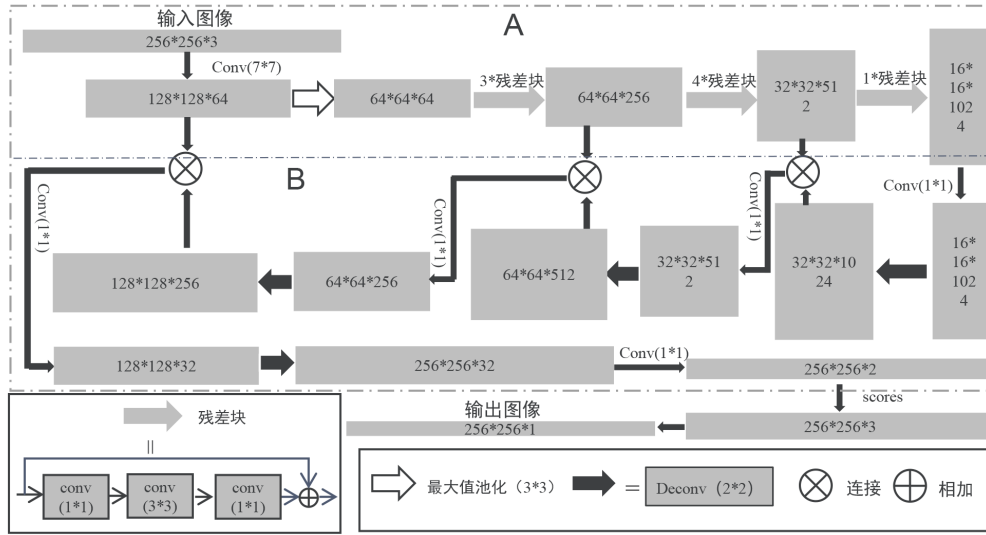


图 5 UNet 网络框架图

Fig.5 The network architecture of UNet

3.2 实验设计

3.2.1 MPSPNet 模型和 UNet 模型耕地提取实验

本研究基于深度学习模型的耕地提取流程如图 6 所示，首先利用构建好的训练样本集对模型进行训练，并利用验证样本集选择模型最优参数，得到最优模型，最后对山东省的所有样本进行耕地提取测试以评估模型性能。

两种模型训练时均使用 Adam 算法(Kingma 和 Ba, 2015)作为梯度下降优化方法。为了加快训练进程，本研究直接迁移在 ImageNet 数据集上预训练好的 ResNet-50 模型进行网络训练。设置每训练 2 块训练样本数据，更新一次模型参数，每 1000 次迭代进行一

次测试，得到一个损失值。为了使得模型的梯度下降法能够有较好的性能，要把学习率的值设定在合适的范围内，学习率决定了参数移动到最优值的速度快慢。如果学习率过大，很可能会越过最优值；如果学习率过小，优化的效率可能过低，长时间算法无法收敛。所以学习率对于算法性能的表现至关重要。本研究参考 Kingma 和 Ba (2015) 中的参数设置，具体参数设置如下：学习率为 0.0001，动量为 0.9，权重衰减为 0.00001，设置每隔 200000 步长，学习率衰减为原来的 0.1。本研究模型共完成 10 期训练 (所有训练样本训练一次即为一期)，训练结束后选择最小的验证损失值作为最优参数测试模型性能。

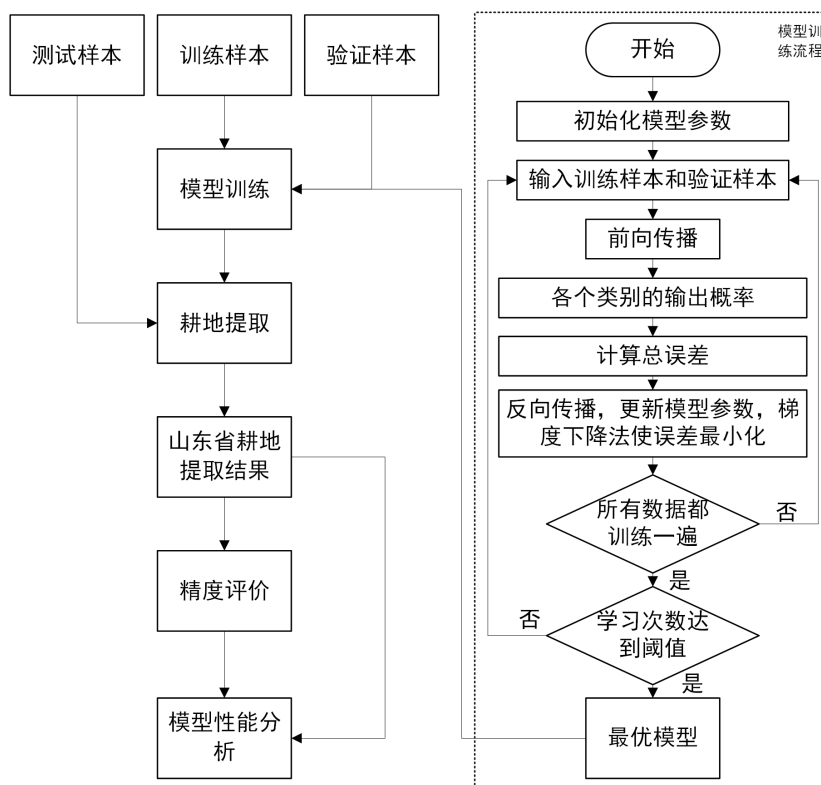


图 6 耕地提取流程图

Fig.6 Flowchart of cropland extraction

3.2.2 与传统面向对象分类方法的耕地提取对比实验

为探究深度学习模型相比于传统方法在高分辨耕地提取上的优势,我们选择面向对象的随机森林图像分析方法(Object-Based Image Analysis-Random Forest, OBIA-RF)与两种深度学习模型进行对比。面向对象的分类方法以特征相近的多个像元为基础分类对象,可以更大程度地利用高分辨影像丰富的特征信息,相比于传统的基于像元的分类,更适用于高分影像的耕地提取(Ma 等, 2017)。随机森林分类器(Random Forest, RF)是一个包含多个决策树的分类器,在处理冗余特征时比其他分类器更具有优势,在基于对象的分类中表现较好(Ma 等, 2017)。

考虑到面向对象的分类方法在全省大范围耕地提取中的复杂性和难度,我们在平原区和混合区各随机选取一个区/县,利用网格作为基本单元,将每个研究

区划分为 9 个子区域, 随机将其划分为训练集、验证集和测试集(图 7), 其中区/县 1 的耕地均为平整规则耕地, 区/县 2 的耕地大部分为破碎不规则耕地, 少部分为平整规则耕地。

面向对象方法的特征选择参考 Ce 等 (2018) 的设置, 将光谱特征(均值和标准差)、纹理特征(灰度共生矩阵)和几何特征(形状指数、周长-面积比)输入随机森林模型进行分类。为比较不同方法的可迁移性和泛化能力, 我们将分别基于区/县 1 和区/县 2 经 OBIA-RF 模型、MPSPNet 模型和 Unet 模型训练得到的模型分别命名为 OB-1、M-1、U-1、OB-2、M-2、U-2, 用训练得到的 6 个模型分别对两个区/县的测试集进行耕地提取测试并进行精度评价。为便于比较分析, 我们将和测试集在同一个区域得到的训练模型称为区域内模型, 和测试集在不同区域得到的训练模型称为区域间模型。

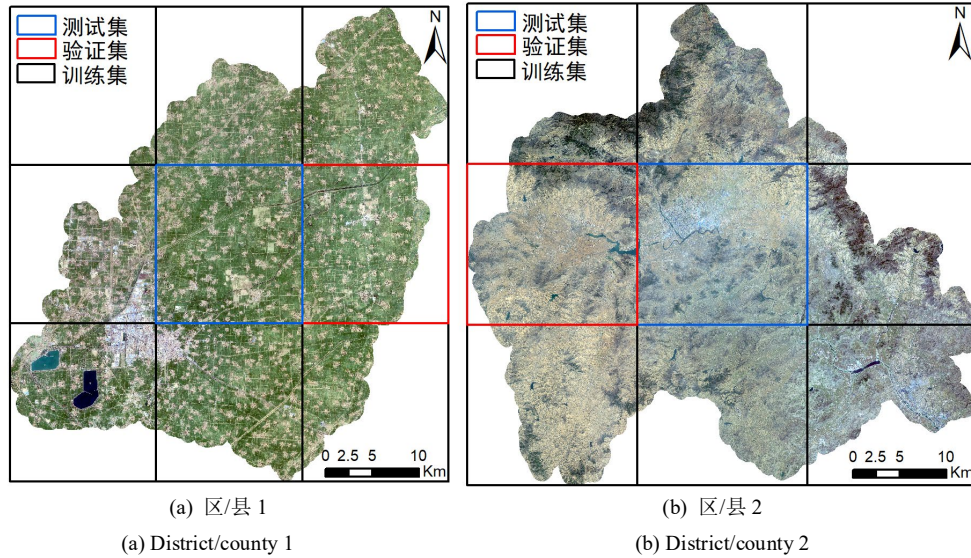


图 7 样本集分布

Fig.7 The distribution of each samples

3.2.3 MPSPNet 模型和 UNet 模型时间迁移实验

为探究两模型的时间迁移能力即模型对年际间特征或类型发生变化的耕地提取效果（Phalke A R 和 Özdoğan M,2018），我们将已经训练好的适用于山东省的 MPSPNet 及 UNet 耕地提取模型分别应用于山东省 2017 年光谱特征及土地覆盖类型发生较大变化的章丘市和苍山县的影像进行耕地提取测试，并与 2016 年/2015 年的耕地提取精度进行对比分析。

3.3 精度评价方法

为定量评价两种模型的分类结果，本研究利用混淆矩阵对提取结果进行全样本精度评价，采用总体分类精度（Overall Accuracy, OA），Kappa 系数，生产者精度（耕地）(Producer's Accuracy ,PA)及用户者精度（耕地）(User's Accuracy ,UA)四种分类精度评价指标。

3.4 不确定性计算

训练完成之后的模型进行耕地提取时，除了输出最终的像元类别结果外，还输出像元属于某一给定类的概率。在分类完成之后，分类结果的正确性可以通过精度评价指标来评估，但是不能反映分类结果的不确定性。可视化和评估分类过程中的不确定性可以用于进一步评估模型的性能表现。为了评估分类结果的不确定性，可以仅考虑像元属于最终类的概率，但这种方法没有考虑到像元类概率的分布。为了在不确定性评价中完整体现像元属于各个类的概率，本研究利用类概率分布的熵计算不确定性（Dehghan 和 Ghassemian, 2006; François 等, 2015），基于熵的归一

化不确定性 (Normalized Uncertainty Criteria, NUC)计算公式如下：

$$NUC = 1 - \frac{\log_2 M - (-\sum_{i=1}^M P_i \times \log_2 P_i)}{\log_2 M} \quad (1)$$

NUC 是一个像元的不确定性，M 是目标分类个数， P_i 是该像元被分为类别 i 的概率，NUC 的范围在 0 到 1 之间。NUC 越高，该像元被识别为最终类的确定性越低，说明模型识别该像元的能力越差。NUC 越低，该像元被识别为最终类的确定性越高，如果此时最终分类结果与真实类别一致，则表明模型正确地识别该像元的概率大，相反，则说明模型误识别该像元的概率大。例如一个像元被模型分为耕地和非耕地的概率都是 50%，则该像元的不确定性为 1，说明模型几乎无法识别该像元。为便于可视化和分析像元的不确定性，本研究将所有的不确定性结果乘以 100 表示。

4 实验结果

4.1 MPSPNet 模型和 UNet 模型耕地提取结果

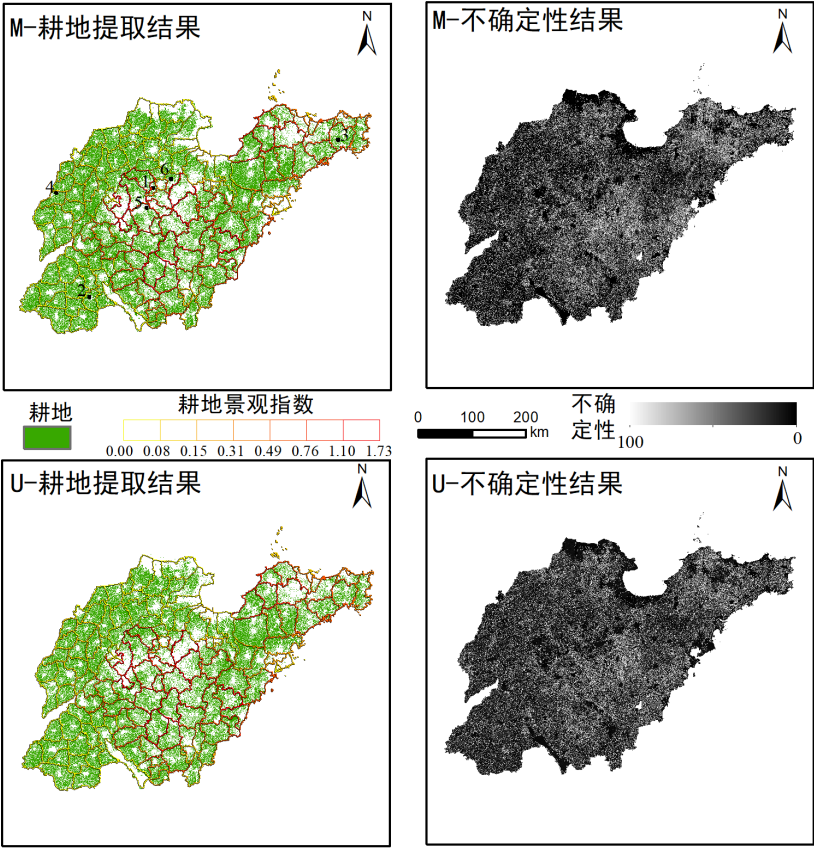
4.1.1 耕地提取结果

两模型的山东省耕地提取结果如图 8 所示，MPSPNet 模型及 UNet 模型的耕地提取结果分别用 M 和 U 表示，图 8 (a) 为两模型山东省耕地的总体提取结果。耕地比例较高的区/县主要分布在平原区域的内陆区域以及混合区域中中部山区以外的内陆区域，耕地比例较低的区/县主要分布在山东省的主城区、中部山区以及沿海区域。为更具体地分析模型耕地提取效

果，我们随机选取了平原区域和混合区域不同时期耕地形态及有少量云覆盖的6块2000x2000像元的影像，如图8(b)所示。

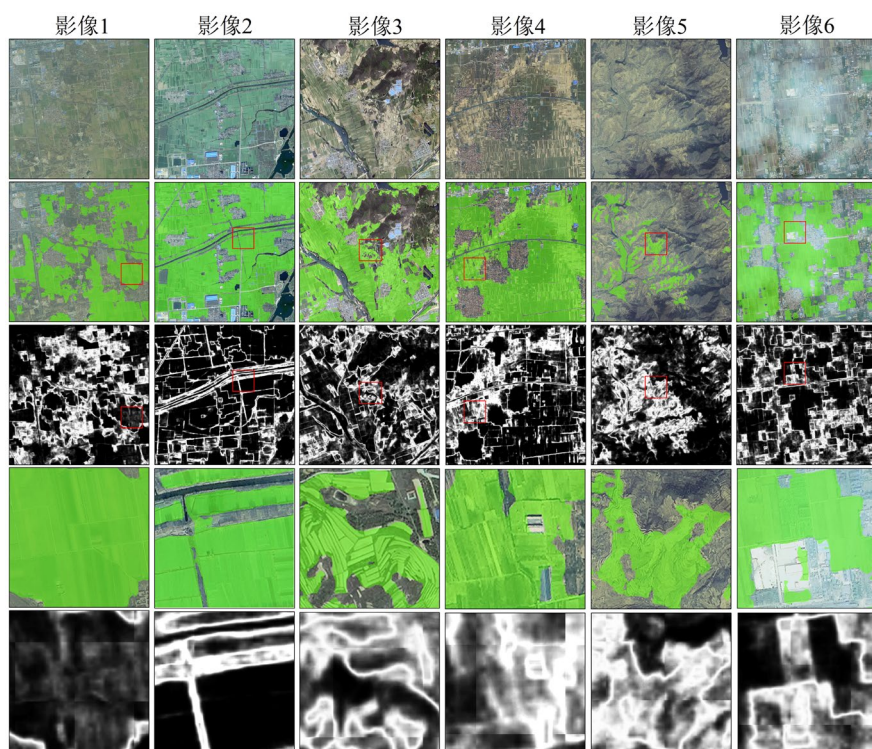
总体来看，两种模型的分类效果无明显差异。在地形平坦、地块平整、耕地面积大的平原区域，两模型能较完整地提取耕地，如图8(b1)、(b2)、(b4)、(c1)、(c2)、(c4)所示。模型对无植被覆盖的耕地、作物生长不同时期的耕地等都能较完整的提取；对于与城市、道路、水体等其他用地相邻近的耕地，模型能较好地识别边界。模型不仅能识别出纹理形状规则、区分度大的平原区域耕地，在混合区形状不规则、破碎、区分度低的耕地，模型也能较为准确的提取，如图8(b3)、

(c3)所示。在地形复杂的区域，模型能较好地地区分出山区林地与耕地，山区形状不规则、分布较零星的耕地大多数能较为准确地提取。在影像质量较差，有少量薄云覆盖的区域如图8(b6)、(c6)所示，两模型也能较为准确的识别。两种模型对耕地的提取仍存在错分和漏分的现象。模型会将少量的滩涂、建设用地等错分为耕地，对于丘陵上的不规则耕地不能完全提取。这些地类与无植被覆盖的耕地特征较为接近，目视难以识别，如图9所示。综上所述，两种模型都能较完整地识别影像中能够目视识别的大部分耕地且无明显差异，对于影像中目视识别有困难的耕地，模型的识别效果较差。



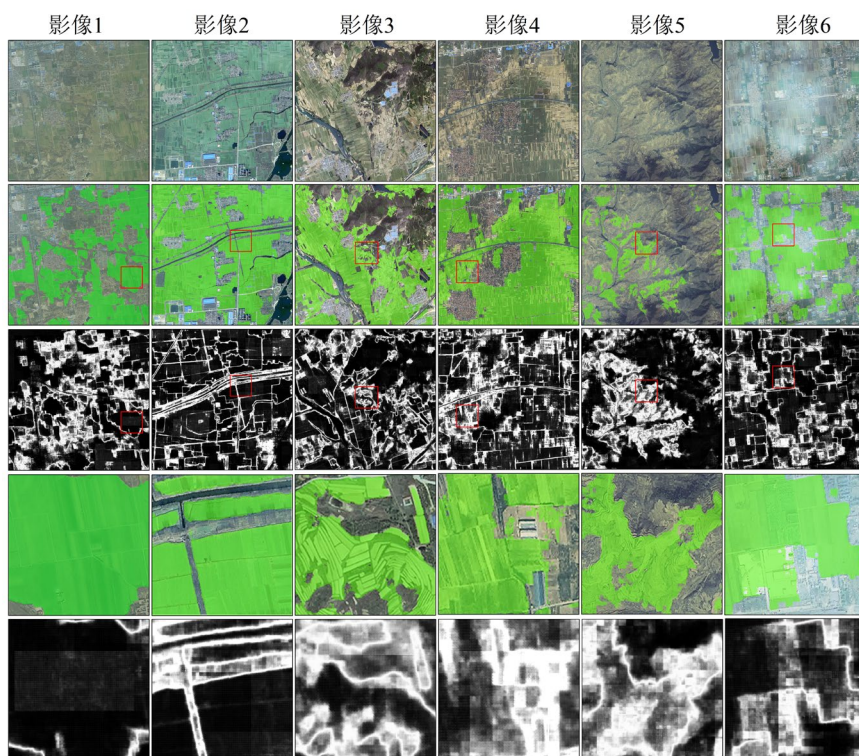
(a) 山东省耕地提取结果及不确定性结果

(a) The results of cropland extraction and uncertainty in Shandong province



(b) MPSPNet 模型局部区域耕地提取结果及不确定性结果

(b) The results of cropland extraction and uncertainty in areas from MPSPNet



(c) UNet 模型局部区域耕地提取结果及不确定性结果

(c) The results of cropland extraction and uncertainty in areas from UNet

图 8 耕地提取结果及不确定性结果

Fig.8 The results of cropland extraction and uncertainty

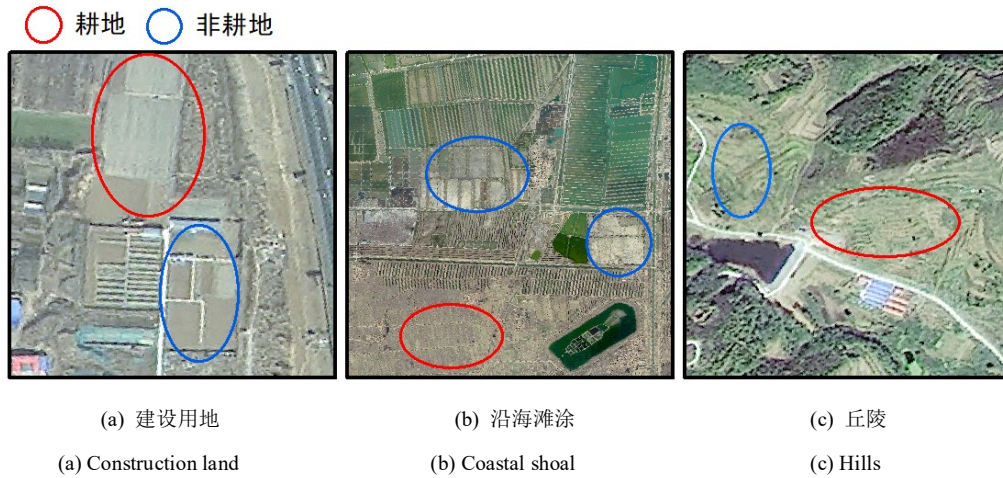


图 9 易被错分的区域

Fig.9 The areas of misclassified easily

4.1.2 精度评价结果

本研究从全省、区域、区/县三种尺度上分别对两种模型的分类结果进行评价。全省及区域尺度评价结果如表 3 所示。整体来看，两模型在各区域的精度评价指标之间无明显差异。

在全省尺度上，两个模型的总体精度都达到了 0.90，Kappa 系数在 0.79 以上，PA 及 UA 的值都在 0.89 以上。在区域尺度上，测试集与全省在各指标之间无明显差异，说明模型对参与训练的样本和没有参与训练的样本的提取效果无明显差异。两模型在平原

区域的各评价指标均高于混合区域，两模型在各评价指标之间差值小于 0.04，MPSPNet 模型在各类区域的 OA 值、Kappa 值、PA 值均略高于 UNet 模型，UNet 模型在各类区域的 UA 值均略高于 MPSPNet 模型。

精度评价结果表明两个模型在全省及区域尺度上的耕地提取效果较好且无明显差异，平原区域的分类精度高于混合区域。两模型在全省和测试集上的提取精度无明显差异，证明两模型具有较强的泛化能力。

表 3 全省及区域精度评价结果

Table 3 The quantitative evaluation results of cropland extraction in province and areas

		OA	Kappa	PA	UA
MPSPNet	全省	0.90	0.80	0.91	0.89
	测试集	0.90	0.79	0.91	0.89
	平原	0.91	0.81	0.94	0.91
	混合	0.89	0.77	0.89	0.88
UNet	全省	0.90	0.79	0.89	0.91
	测试集	0.89	0.78	0.88	0.90
	平原	0.90	0.80	0.91	0.92
	混合	0.88	0.75	0.85	0.89

山东省各区/县的精度评价结果如图 10 所示，整体来看，两模型在非测试集区/县的评价指标趋势分布与测试集区/县的评价指标趋势分布相似，表明模型在区/县尺度上对测试集和非测试集的耕地提取精度无明显差异。

两模型在所有评价指标之间的差值小于 0.05，

UNet 模型和 MPSPNet 模型分别有 61%和 69%的区/县 OA 值高于 0.90。92%的区/县 Kappa 系数大于 0.60，PA 及 UA 高于 0.60 的区/县比例分别为 93%和 97%。两模型在各区/县四个评价指标之间无明显差异，部分区/县的评价指标较低，表明两模型的选择对各区/县耕地提取效果没有影响，少数区/县存在较严重的错分及

漏分。本研究分别选取了平原区域及混合区域中评价指标较低的 8 个区/县（如图 10 所示）具体探究耕地错分及漏分情况。区/县 4 分布在丘陵区域，耕地多为裸露状态，形态不规则，较为破碎，目视不易识别，漏分较多。部分丘陵区草地、林地、以及部分浅海滩涂海产养殖区与耕地纹理特征相似（图 8），目视难以区分。区/县 1 为主城区，极少量的耕地在城区中难以识别。其余区/县位于平原区，耕地景观指数较低，耕地较多且较为平整，但城区邻近无作物覆盖的耕地、河流沿岸的滩涂不规则耕地与城区内建筑区域的裸地相近，目视难以区分，如图 9 所示。综上所述，两模型在山东省各区/县的耕地识别精度无明显差异，在目视难以识别的区域，存在较严重的耕地错分及漏分情况。

为定量分析模型对山东省耕地的提取效果，本研究根据山东省的地势及耕地形态分布情况，以区/县

为单位，基于第三次全国农业普查二类测量成果转化后的 2m 分辨率栅格数据，采用景观指数中的边缘面积比（Perimeter Area Ratio，PARA）及区/县高程标准差计算各区/县的耕地景观形态特征指标：

$$SD_PARAi = SDEMi \times Edgei_{crop} / Totali_{crop} \quad (2)$$

SD_PARAi 是 i 区/县的耕地景观指数， $SDEMi$ 是 i 区/县的高程标准差， $Edgei_{crop}$ 是 i 区/县的耕地边缘像元数， $Totali_{crop}$ 是 i 区/县的耕地总像元数。耕地景观指数越高表明该区/县的耕地形态越复杂。

两模型在测试集 119 个区/县的 OA 值和 Kappa 系数及山东省耕地景观指数的空间可视化结果如图 11 所示。两模型在各指标之间的分布特征无明显差异，耕地景观指数与 OA 值（MPSPNet 精度与 UNet 精度之间无明显差异，这里以 MPSPNet 精度为例进行分析。）在空间分布特征上存在一定规律。

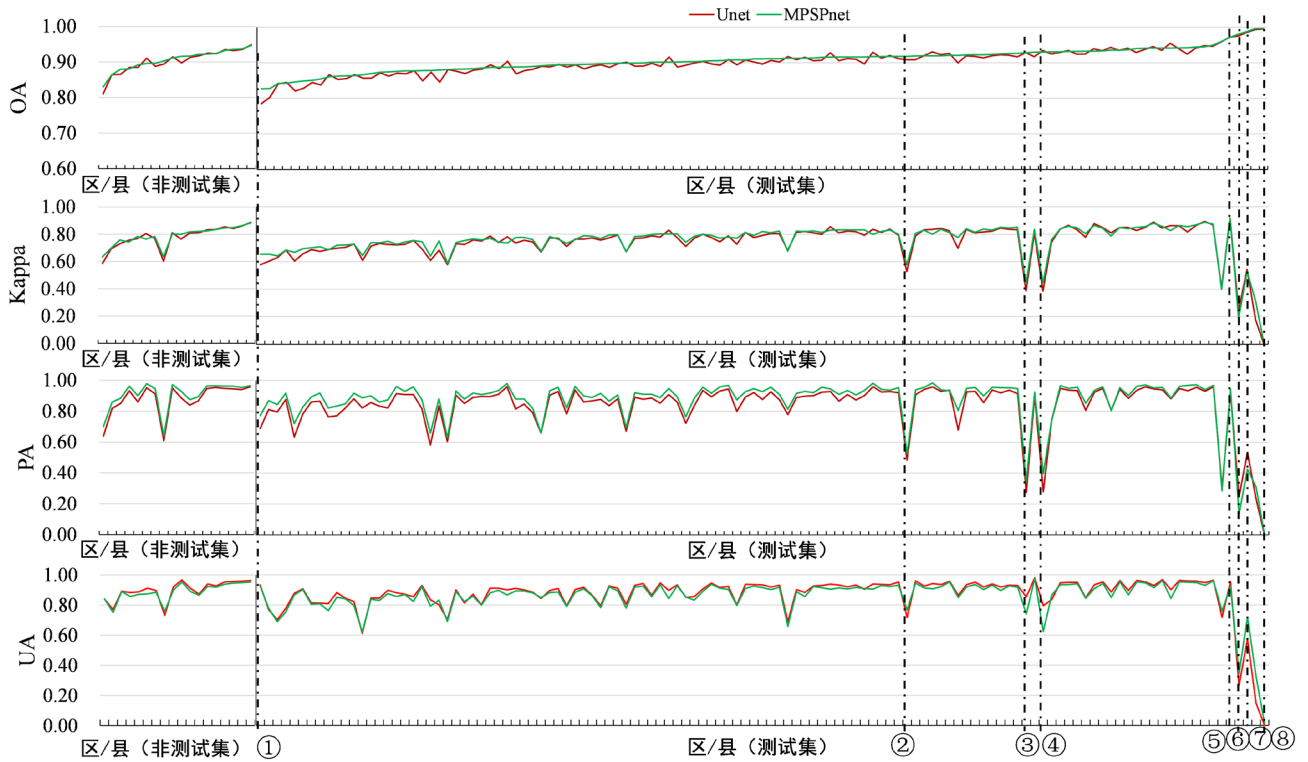


图 10 区/县精度评价结果

Fig.10 The results of quantitative evaluation in district or county

如表 4 所示，耕地景观指数在 0.15 以下的 60 个区/县中 72% 的区/县 OA 值高于 0.90，这些区/县主要分布在平原区的大部分区域和混合区域中邻近平原区的北部区域及最南部区域。其余 28% 的区/县 OA 值范围在 0.83-0.89，主要分布在平原区域北部沿海区域，该区域存在大量与耕地纹理特征相似的滩涂海产养殖区。耕地景观指数在 0.15-0.75 的 37 个区/县中 76% 的

区/县 OA 值在 0.90 以上，这些区/县主要分布在混合区邻近中部丘陵的区域、东北部的沿海区域以及南部与平原区相邻的区域。耕地景观指数在 0.75 以上的共有 22 个区/县，其中 9 个区/县 OA 值高于 0.90，分布在邻近平原区的区域，其余区/县精度在 0.86-0.89 之间，主要集中在混合区的中部丘陵区，耕地形态更为复杂，目视也不易识别。

以上结果表明精度较高的区/县主要分布在大部分平原区域以及混合区域中部分耕地景观指数较低的区域。精度较低的区/县主要分布在平原区北部沿海区域和混合区耕地景观指数较高的中部丘陵区域以及沿海区域。

表 4 整体精度在耕地景观指数的分区统计

Table 4 Regional statistics of OA in cropland landscape index

耕地景观指数 数区间	区/县个数	OA 值范围	OA 平均值	区/县个数 (OA 值>0.90)
0.00-0.15	60	0.83-0.99	0.91	43
0.15-0.75	37	0.84-0.94	0.90	28
0.75-1.73	22	0.86-0.93	0.89	9

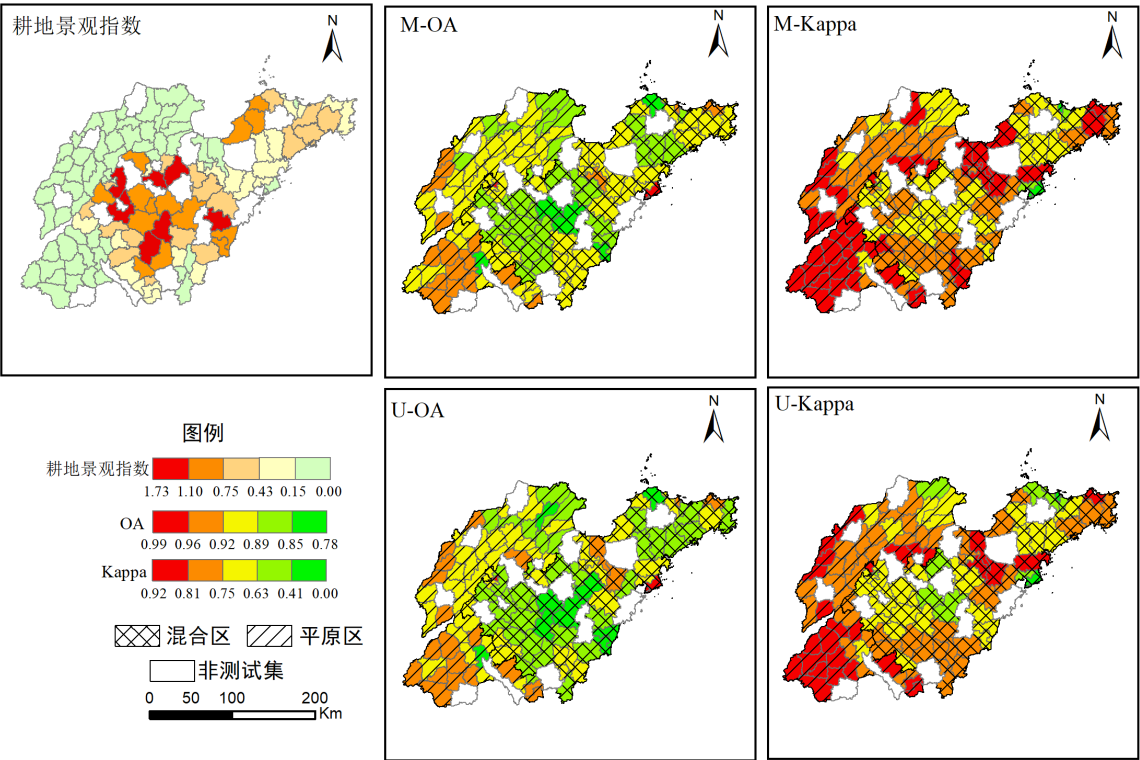


图 11 测试集耕地景观指数和精度评价结果空间分布

Fig.10 Landscape index of cropland and spatial distribution of quantitative evaluation in district or county

4.1.3 不确定性结果

根据 NUC（公式 1）计算得到的两种模型的不确定性结果如图 8（a）、图 8（b）及 8（c）所示，不确定性较高的区域集中在平原区域的北部近海区域以及混合区域的中东部及近海区域（图 8（a））。在较为平整，纹理较为简单的耕地田块（图 8（b1）、（c1））以及非耕地中的中心城区（图 8（b2）、（c2））、水域（图 8（b3）、（c3））、山区非耕地（图 8（b5）、（c5））等区分度较大的区域不确定性较低，此类区域的纹理特征较为简单，模型的识别效果也较好。在

耕地与非耕地的临界区域（图 8（b2）、（b3）、（b6）、（c2）、（c3）、（c6））、纹理特征较为复杂的丘陵区域耕地（图 8（b5）、（c5））以及各地类的边界区域不确定性较高（图 8（b）、（c）），模型对该区域的耕地及非耕地识别效果较差。

为定量探究模型对影像中耕地特征的学习效果，本研究选取了具有代表性的不同耕地景观形态及不同评价精度的 3 个区/县，对两模型耕地识别的不确定性进行分析统计，如图 12 所示。模型对像元的识别结果可以分为正确分类和错误分类，正确分类是指耕地被正确分为耕地，耕地错误分类是指耕地被误分为非耕

地。对于正确分类像元，其不确定性越低，表明模型认为该像元有很大概率属于正确的类别，表明模型对耕地的特征学习效果及识别效果较好。对于错误分类像元，其不确定性越高，表明模型认为该像元属于耕地的概率和非耕地的概率相近，分类结果改善的可能性越大。正确分类像元的不确定性越低，错误分类像元的不确定性越高，表明模型对耕地特征的学习效果越好，分类性能更好。

在内陆平原区域的金乡县，两模型的不确定性分布相似。在正确分类的像元中，两模型不确定性在 20 以下的像元比例分别为 83%和 86%，在错误分类的像元中，两模型不确定性在 80 以上的像元比例分别为 51%和 55%。该结果表明两模型在此区域有很大把握正确识别大多数像元，错误分类的像元多数不确定性较高，说明两模型认为多数分类错误的像元属于耕地和非耕地的概率相近。

在丘陵区域的黄岛区，两模型的不确定性分布存在差异。在正确分类的像元中，MPSPNet 模型有 64%的像元的不确定性在 40 以下，UNet 模型有 69%的像元的不确定性在 40 以下。在错误分类的像元中，MPSPNet 模型和 UNet 模型分别有 75%的像元和 57%的像元的不确定性在 60 以上。该结果表明两模型有较大把握正确识别多数像元，MPSPNet 模型分类错误的像元中多数不确定性较高，即模型认为该类像元属于

耕地和非耕地的概率接近。UNet 模型与 MPSPNet 模型相比，有更多分类错误的像元不确定性较低，表明模型有较高的把握认为该类真实类别为耕地的像元属于非耕地。

在沿海区域的河口区，两模型的不确定性分布存在较大差异。在正确分类的像元中，MPSPNet 模型和 UNet 模型分别有 63%和 44%的像元不确定性在 20 以下。在错误分类的像元中，MPSPNet 模型有 59%的像元的不确定性在 60 以上，UNet 模型仅有 19%的像元不确定性在 60 以上。该结果表明 MPSPNet 模型能有较高把握正确识别多数像元，识别错误的像元多数不确定性较高。UNet 模型多数识别错误的像元的不确定性较低，表明模型认为这些真实类别为耕地的像元属于非耕地的概率更大。

根据以上分析，两模型对耕地特征的学习效果在不同的区域类型中存在差异。在耕地特征更为简单的区域，两模型对耕地特征的学习及识别效果较好，无明显差异。在耕地及非耕地结构较为复杂的区域，两模型的识别效果存在差异。在耕地形态不规则，较为破碎的区域，耕地的特征更为复杂，两模型对此类耕地特征的学习效果均较差。对于存在与耕地特征纹理特征较为相近地物的区域，模型对于耕地的识别效果一般，且 UNet 模型有更大的概率将耕地误分。

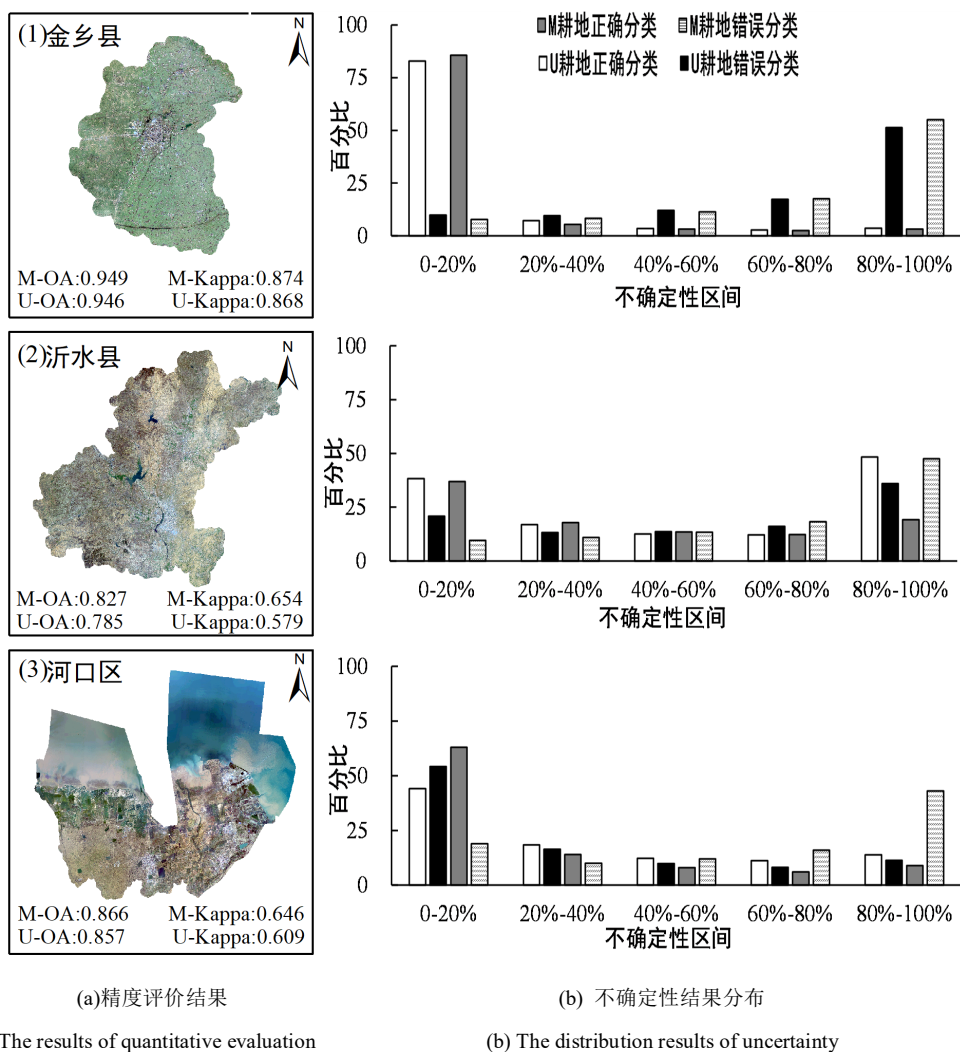


图 12 区/县不确定性分析结果

Fig.12 Results of uncertainty analysis in district or county

4.2 与传统面向对象分类方法的对比实验结果

从耕地提取精度（表 5）和耕地提取结果（图 13）来看，MPSPNet 模型和 UNet 模型在两个测试集上的表现都优于 OBIA-RF 模型。

从区域内模型的耕地提取效果看，在测试集 1 中，M-1 和 U-1 的耕地提取精度高于 OB-1，M-1 和 U-1 的耕地提取结果与真实耕地无明显差异，OB-1 的耕地提取结果存在部分漏分及错分，如图 13（ac）所示。在测试集 2 中，U-2 的耕地提取精度最高，OB-2 的提取精度最低，U-2 的耕地提取结果与真实耕地较为接近，OB-2 模型存在较大的漏分及误分。从区域间模型

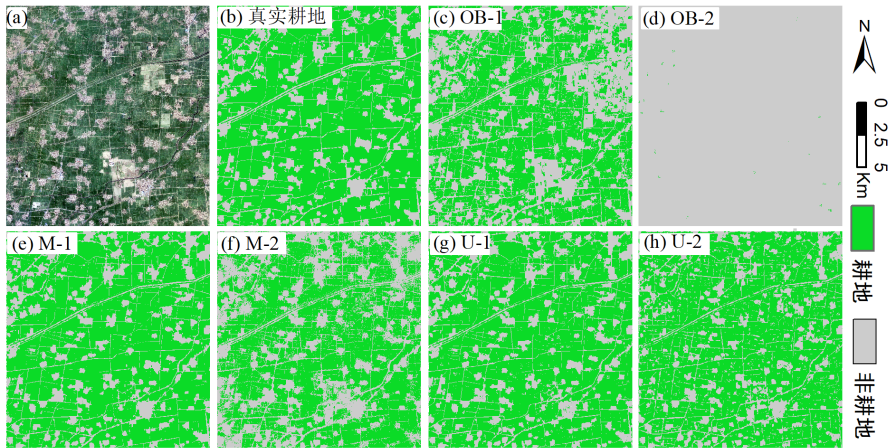
看，在测试集 1 中，M-2 和 U-2 的耕地提取精度高于 OB-2，OB-2 模型几乎无法识别该区域的耕地，如图 13（ad）所示。在测试集 2 中，OB-1、M-1 和 U-1 三个模型的提取精度无明显差异，OB-1 模型精度略低，3 个模型耕地提取结果与真实耕地都存在较大差异。

综上所述，无论是区域内模型还是区域间模型，三类模型对测试集 1 的提取精度都高于测试集 2，在区/县 1 训练得到的模型在测试集 2 上的表现都较差。总体来看，MPSPNet 模型和 UNet 模型在平原区域的测试集 1 和丘陵区域的测试集 2 的耕地提取效果、迁移能力和泛化能力都优于传统的 OBIA-RF 分类方法。

表 5 MPSPNet 模型、UNet 模型及 OBIA-RF 模型在两测试集的精度评价结果

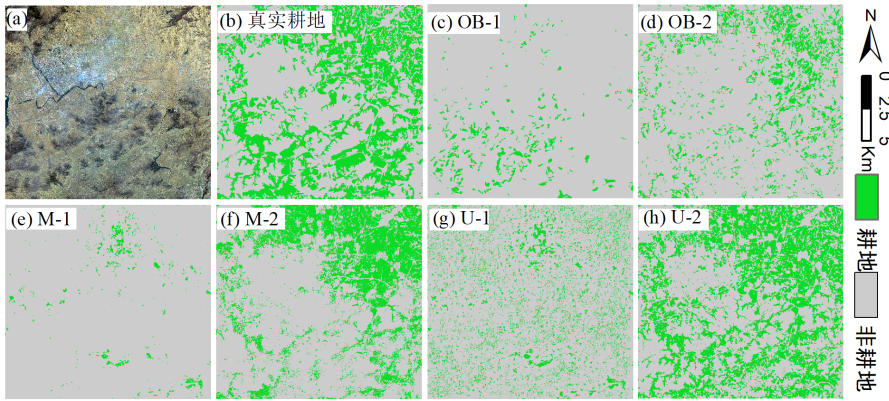
Table 5 The quantitative evaluation results in two test samples from MPSPNet, UNet and OBIA-RF

		OB-1	OB-2	M-1	M-2	U-1	U-2
区/县 1	OA	0.81	0.25	0.94	0.83	0.94	0.87
	Kappa	0.57	0.00	0.84	0.61	0.83	0.66
	PA	0.76	0.00	0.96	0.79	0.95	0.91
	UA	0.97	0.27	0.96	0.97	0.96	0.92
区/县 2	OA	0.61	0.66	0.64	0.77	0.62	0.83
	Kappa	0.00	0.16	0.04	0.47	0.05	0.65
	PA	0.04	0.23	0.04	0.56	0.14	0.80
	UA	0.37	0.61	0.80	0.75	0.47	0.76



(a) 测试集 1 耕地提取结果

(a) The results of cropland extraction in test samples 1



(b) 测试集 2 耕地提取结果

(b) The results of cropland extraction in test samples 2

图 13 MPSPNet 模型、UNet 模型及 OBIA-RF 模型在两测试集的耕地提取结果

Fig.13 The results of classification in two test samples from MPSPNet ,UNet and OBIA-RF

4.3 MPSPNet 模型和 UNet 模型时间迁移实验结果

两区/县耕地景观指数分别为 1.41、0.61，提取结果如图 14 所示，两区/县的时间及精度评价结果如表 4 所示。无论是耕地生长状态不同引起的耕地特征的变化还是土地类型的变化（耕地变为非耕地或者非耕

地变为耕地），两模型都可以很好地识别耕地与非耕地。从精度评价结果来看，两模型之间无明显差异，2017 年的总体精度比 2016 年/2015 年的总体精度降低 0.05 左右。结果表明两模型在影像时相特征发生变化时，仍能准确识别耕地，证明了两耕地提取模型的时间可迁移性。



(a) 章丘市耕地提取结果

(a) The results of cropland extraction in Zhangqiu City



(b) 苍山县耕地提取结果

(b) The results of cropland extraction in Cangshan County

图 14 基于时间迁移的模型分类结果

Fig.14 The classification results of models based on time migration

表 6 区/县时间及精度评价结果

Table 5 The time of district or county and the quantitative evaluation results

区/县	2016/2015 年时间	2017 年时间	2016 年/2015 年精度		2017 年精度	
			OA-M	OA-U	OA-M	OA-U
章丘市	20160224	20170530	0.93	0.94	0.88	0.87
苍山县	20150410	20170528	0.92	0.92	0.87	0.86

5 讨论

两种模型在全省及各区域尺度上的耕地提取效果较好且无明显差异，同时在时间上迁移应用也取得较好的精度，表明这两种方法的选择不会影响耕地的提取效果。原因主要有两点：一是参与训练的模型样本量足够大，且样本中的耕地时相几乎涵盖全省影像中耕地的各个形态（图 3），模型在训练时能够学习到各个耕地形态的特征，在测试时能够依据这些特征识别耕地。二是两个模型中的多层卷积层的设计可以挖掘影像中的深层次和浅层次特征，模型在训练完后学习到了耕地和非耕地更本质的特征，有助于准确识别耕地。但是从不确定性分析来看，在耕地形态不规则的区域以及非耕地特征与耕地特征不易区分的区域，

UNet 模型对耕地及非耕地的特征学习效果不如 MPSPNet 模型，UNet 更容易将非耕地错误识别为耕地。这主要是由于两模型虽然都能够通过卷积层学习到影像中的深层和浅层次特征，但是 MPSPNet 中的金字塔池化模块可以学习到包含不同尺度、不同子区域间关系的分层全局特征信息，对于影像特征的信息学习更为全面，有助于耕地的识别（Zhang 等，2020），而 UNet 模型中缺乏此类信息。

耕地景观指数作为反映耕地景观特征的参数，与深度学习模型耕地提取效果间具有一定的关联。两种模型可较好地识别出耕地景观指数较低的内陆平原区域的耕地，而丘陵区域较为破碎的耕地景观指数较高的耕地易被漏分，模型的识别效果较差。表明耕地的景观形态特征是影响模型耕地提取效果的重要因素。此外模型在平原区域的滩涂和建设用地区域容易出现

较严重的耕地错分问题,主要是因为平原区域的滩涂、部分建设用地等与无植被覆盖的耕地纹理特征相近,目视无法准确识别,模型学习到的此类地物的特征与裸露耕地的特征相近,模型几乎无法判断,造成了耕地错分。另外,不同于模型容易学习的景观特征简单的平整耕地地块,分布在山脚或山坡处的耕地景观复杂,形状破碎,特征复杂,模型学习这类特征的难度较大,并且这些耕地多数与山区的裸地等纹理相近,造成了耕地漏分。

相较于传统的 OBIA-RF 方法,MPSPNet 模型和 UNet 模型能取得更好的分类效果且具有更强的特征学习能力和泛化能力。无论是在耕地平整规则区域还是耕地破碎不规则区域,三种方法应用在同类区域的耕地提取时,两种卷积神经网络的耕地提取效果都优于 OBIA-RF 方法,表明 MPSPNet 模型和 UNet 模型的特征学习效果更好,主要是由于两种模型的网络架构根据训练样本可以学习到影像中不同的深层次和浅层次特征,包括不同形态耕地的不同尺度的光谱特征、纹理特征、边界特征以及上下文信息等,预测时能够根据这些耕地及非耕地的更本质的特征准确识别出各类型耕地,有效地避免临近地物的干扰。三种方法在平原区/县训练得到的模型应用在混合区/县时,耕地提取效果均较差且无明显差异,这主要是由于训练样本中缺少破碎不规则耕地,模型无法学习到此类耕地特征,不能识别混合区域中大部分破碎不规则耕地。三种方法在混合区/县训练得到的模型应用在平原区/县时,OBIA-RF 方法耕地提取效果最差,这主要是由于两种卷积神经网络能利用训练样本中少量的平整规则耕地,学习此类耕地特征,应用于平原区域的耕地提取,而传统 OBIA-RF 方法所采用的先分割再进行分类的过程,不利于挖掘影像中地物不同尺度下更为深刻和本质的特征(Cheng 等, 2017),而且这种方法的分类精度很大程度上受分割效果的影响,在高分辨率遥感影像中过分割和欠分割问题的存在,导致模型分类效果较差(Shu 等, 2015)。

本研究探究了两模型在大范围高精度耕地提取上的适用情况,全省整体精度优于 90%,相比于传统方法,耕地提取效率显著提升,在节约生产成本的同时,满足了耕地产品的实用精度需求,给未来耕地实时监测和保护等提供了一种切实可行的新思路。虽然两种模型在大尺度耕地提取上的精度较高,但是在耕地景观特征复杂以及耕地与非耕地特征相近的区域,模型仍存在严重的错分及漏分问题。因此未来对于高分辨耕地的自动化提取可以通过改进模型对于此类区域细

节特征的挖掘和学习来提高模型对此类耕地的识别效果。另外,本研究发现耕地景观指数与模型的耕地提取精度之间存在一定的关联,未来的研究可以考虑依据耕地景观指数来改善训练样本中耕地特征复杂区域的比例以提高模型对此类区域特征的学习。

6 结论

本文以山东省为研究区,以区/县为基本单元,利用基于 2015 年至 2017 年的 GF-1 号及 GF-2 号的 RGB 影像镶嵌的覆盖山东省的影像和全国农业普查二类测量成果构建了全省的标记样本数据集。通过构建 MPSPNet 及 UNet 深度学习网络模型并进行训练,得到两个最优网络模型,分别对研究区进行耕地提取测试及精度评价,探究两模型在大范围高精度自动化耕地提取的潜力,研究获得以下主要结论:

(1)在区/县尺度耕地提取上,两种深度卷积神经网络模型比传统的面向对象的分类方法总体精度更高。将 MPSPNet 模型和 UNet 模型应用到全省尺度的耕地提取上,两模型之间的总体精度无明显差异,优于 90%。该研究结果证明了将两种模型应用于全省大范围高分辨的耕地自动化提取是可行的,解决了传统分类方法中训练样本的“特定区域、特定影像、特定目标”的局限性,分类结果具有较高的实际应用价值。

(2)模型的选择对同一区域的耕地提取精度无明显影响,在不同的区域,模型耕地提取精度与区域的耕地景观特征有关。平原区域的耕地提取总体精度比混合区域总体精度高 0.02,在目视容易判断的地块平整、规则分布的耕地景观指数较低的区域模型提取效果较好,平均总体精度为 0.90。在耕地破碎,不规则分布的耕地景观指数较高的丘陵区域以及沿海滩涂、建设用地等与裸露耕地特征相近的区域模型提取效果较差,平均总体精度为 0.89。有待改善,并且在此类区域,UNet 模型更容易将耕地误分。该研究结果证明了区域的耕地景观特征是影响耕地提取效果的重要因素,并且金字塔池化模块学习到的不同区域的上下文信息有助于降低模型的错分概率。

(3)两模型比传统面向对象的方法具有更强的泛化能力和时空迁移能力。两模型在测试集的提取精度与全省的提取精度无明显差异,在不同区域、不同时相的影像中也能得到较好的耕地提取效果,进一步验证了深度学习模型在大范围高精度耕地自动化提取上的潜力。

此外本研究还有需要进一步改进的地方,两模型对于耕地特征较为复杂的区域识别效果较差,可能是由于样本选择时此类耕地的样本不充足导致其学习不

充分,未来可以考虑依据耕地景观指数进一步增强和平衡各类样本,提高模型的分类精度。

参考文献(References)

- Blaschke T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1): 2-16.
- Cheng G, Han J and Lu X. 2017. Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, 105(10):1865-1883.
- Clayton D G. 1971. Gram-Schmidt orthogonalization. *Applied Statistics*, 20 (3) :335 ~ 338.
- China National Bureau of Statistics, Office of the Leading Group for the Third National Agricultural Census of the State Council. 2016. Working plan for remote sensing measurement of crop area in the third national agricultural census. Beijing: China Statistics Press:73-78. (中华人民共和国国家统计局, 国务院第三次全国农业普查领导小组办公室. 2016. 第三次全国农业普查农作物面积遥感测量工作方案. 北京: 中国统计出版社: 73-78)
- Ce Zhang, Isabel Sargent, Xin Pan, Huapeng Li, Andy Gardiner, Jonathon Hare, Peter M. Atkinson. 2018. An object-based convolutional neural network (OCNN) for urban land use classification, *Remote Sensing of Environment*, 216:57-70 [DOI:https://doi.org/10.1016/j.rse.2018.06.034]
- Chen L, Zhu Y, Papandreou G, Schroff F and Adam H. 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *ArXiv*, abs/1802.02611.
- De Pinho C M D, Fonseca L M G, Korting T S, et al. 2012. Land-cover classification of an intra-urban environment using high-resolution images and object-based image analysis. *International Journal of Remote Sensing*, 33 (19) : 5973-5995.
- Dehghan H. and Ghassemian H. 2006. Measurement of uncertainty by the entropy: application to the classification of MSS data. *International Journal of Remote Sensing*, 27(18), 4005–4014. [DOI:10.1080/01431160600647225]
- François W, Guadalupe S C and Pierre D. 2015. Automated annual cropland mapping using knowledge-based temporal features, *ISPRS Journal of Photogrammetry and Remote Sensing*, 110:1-13. [DOI:10.1016/j.isprsjprs.2015.09.013.]
- Goodin D G, Anibas K L, Bezymennyi M. 2015. Mapping land cover and land use from object-based classification: An example from a complex agricultural landscape. *International Journal of Remote Sensing*, 36(18): 4702-4723.
- Hu Q, Wu W B, Song Q, Yu Q 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.05781752.2015.10.004])
- Hernandez I E R, Shi W Z. 2018. A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. *International Journal of Remote Sensing*, 39(4):1175-1198.
- He K, Zhang X, Ren S, et al. 2016. Deep Residual Learning for Image Recognition. *Computer Vision and Pattern Recognition*, Las Vegas, 770-778.
- Hugo Costa, Giles M. Foody, Doreen S. Boyd. 2017. Using mixed objects in the training of object-based image classifications. *Remote Sensing of Environment*, 190:188–197.
- Julien Y, Sobrino J A, Jiménez-Muñoz J C. 2011. Land use classification from multitemporal Landsat imagery using the Yearly Land Cover Dynamics (YLCD) method. *International Journal of Applied Earth Observation and Geoinformation*, 13(5):711-720.
- Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R and Darrell T. 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. *arXiv: Computer Vision and Pattern Recognition*.
- Krizhevsky A, Sutskever, I and Hinton G E. 2012. ImageNet Classification with Deep Convolutional Neural Networks. *International Conference on Neural Information Processing Systems*, 1097-1105.
- Kingma D P and Ba J. 2015. Adam: A Method for Stochastic Optimization. *international conference on learning representations*.
- Lu H, Fu X, Liu C, Li L, He Y and Li N. 2017. Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning. *Journal of Mountain Science*, 14(4):731-741 [DOI: 10.1007/s11629 - 016-3950-2]
- Li P, Yu H, Wang P and Li K Y. 2017. Comparison and Analysis of Agricultural Information Extraction Methods Based upon GF2 Satellite Images. *Bulletin of Surveying and Mapping*, 1:48-52 (李鹏, 虞虎, 王鹏等. 2017. 基于 GF2 号卫星影像的农业信息提取方法对比分析. *测绘通报*, 1:48-52.)
- Li C J, Huang H, Li W. 2018. Research on agricultural remote sensing image cultivated land extraction technology based on support vector. *Instrumentation Technology*, 11:5-8. (李昌俊, 黄河, 李伟. 基于支持向量机的农业遥感图像耕地提取技术研究. *仪表技术*, 2018(11):5-8.)
- Lu S, Oki K, Shimizu Y et al. 2007. Comparison between several feature extraction /classification methods for mapping complicated agricultural land use patches using airborne hyperspectral data. *International Journal of Remote Sensing*, 28(5):963-84.

- LeCun Y, Boser B, Denker J S, et al.1989. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4): 541-551.
- Larochelle H, Bengio Y, Louradour J, et al. 2009. Exploring Strategies for Training Deep NeuralNetworks. Journal of Machine Learning Research, 1(10):1-40.
- LeCun Y, Bottou L, Bengio Y, et al.1998.Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11): 2278-2324.
- Long J, Shelhamer E, Darrell T.2015.Fully convolutional networks for semantic segmentation.Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,3431-3440.
- Laben C A, BROWER B V. 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening:US, US6011875.
- Ma L, Liu Y, Zhang X L, Ye Y X, Yin G F and Johnson B. 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152:166-177[DOI: 10.1016/j.isprsjprs.2019.04.015]
- Ma L, Li M C, Ma X X, Cheng L, Du P J, Liu Y X. 2017. A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing.
- Phalke A R, Özdoğan M. 2018. Large area cropland extent mapping with Landsat data and a generalized classifier. Remote Sens. Environ. 219:180-195.
- Ronneberger O, Fischer P, Brox T, et al. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. medical image computing and computer assisted intervention, 234-241.
- Szegedy C, Liu W, Jia Y, et al.2015.Going Deeper with Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, 7-12 June 2015, 1-9.
- Simonyan, K. and Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition.
- Shu Y, Tang H, Li J, et al. 2015.Object-based unsupervised classification of VHR panchromatic satellite images by combining the HDP and IBP on multiple scenes. IEEE Transactions on Geoscience and Remote Sensing,53: 6148-6162.
- Wang Q.2021. Progress of environmental remote sensing monitoring technology in China and some related frontier issues. Journal of Remote Sensing, 25(1):25-36 (王桥.2021. 中国环境遥感监测技术进展及若干前沿问题.遥感学报,25(1):25-36)
- Wang W H, Xia L G, Luo J C, et al.2011. An iterative approach to object-oriented classification of remotely sensed image[J]. Geomatics and Information Science of Wuhan University,36 (10) :1154-1158.(王卫红, 夏列钢, 骆剑承等. 2011.面向对象的遥感影像多层次迭代分类方法研究.武汉大学学报(信息科学版) ,36(10):1154-1158.)
- Xu W N.2020.Research on Cultivated Land Extraction Technology of Remote Sensing Images Based on High Resolution Full Convolutional Network. Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences. (徐文娜.2020.基于高分辨率全卷积网络的遥感影像耕地提取方法研究.中国科学院大学(中国科学院深圳先进技术研究院))
- Zhang M and Huang S Y.2019. Remote Sensing Image Classification Based on Landsat-8.Geomatics&Spatial Information Technology, 42(01):177-180 (张明,黄双燕. 2019.基于 Landsat-8 的遥感影像分类研究.测绘与空间地理信息,42(01):177-180.)
- Zhi He, Han Liu, Yiwen Wang, et al. Generative Adversarial Networks-Based Semi-Supervised Learning for Hyperspectral Image Classification. remote sensing, 2017, 9(1042): 1-27.
- Zhang F, Zhao Z G, Li G, et al. 2019.Study on classification and extraction of agricultural land in Qitai County of Xinjiang based on different classifiers. Xinjiang Agricultural Sciences, 56 (8) :1560-1568. (张峰, 赵忠国, 李刚等. 2019. 基于不同分类器的农用地分类提取.新疆农业科学, 56 (8) :1560-1568.)
- Zhou P C, Cheng G, Yao X W, Han J W.2021. Machine learning paradigms in high-resolution remote sensing image interpretation. Journal of Remote Sensing, 25(1):182-197 (周培诚,程堪,姚西文,韩军伟.2021.高分辨率遥感影像解译中的机器学习范式.遥感学报,25(1):182-197)
- Zhang D J, Pan Y Z, Zhang J S, Hu T G, Zhao J H, Li N and Chen Q. 2020. A generalized approach based on convolutional neural networks for large area cropland mapping at very high resolution. Remote Sensing of Environment, 247,111912.
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid Scene Parsing Network. computer vision and pattern recognition, 6230-6239[DOI: 10.1109/CVPR.2017.660]

High-resolution cropland extraction in Shandong province using MPSPNet and UNet network

LI Qiannan^{1,3}, ZHANG Dujuan^{1,3}, PAN Yaozhong^{*1,2,3} DAI Jiajia^{1,3}

1. State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China;

2. Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China;

3. Institute of Remote Sensing Science and Engineering, Faculty of Geographical Sciences, Beijing Normal University, Beijing 100875, China

Abstract: The rapid development of remote sensing image technology enables a large number of high-resolution remote sensing images to provide good data support for the accurate extraction of cropland and other ground features. However, high-resolution remote sensing images have large data volume and complex features, the artificial visual interpretation and traditional classification methods have limited extraction capabilities which cannot realized large-scale high-precision cropland extraction automatically. Deep learning technology has shown superior performance in the automatic extraction of remote sensing image information due to its strong ability to express features, providing a new idea for the automatic extraction of large-scale cropland. Exploring the application of different typical network models in the extraction of cropland with different landscape features is of great significance to the improvement of the quality and efficiency of cropland extraction. Based on above, the study uses the 2m resolution data fused with GF-1 and GF-2 in 2015-2017 as the data source. Using Modified Pyramid Scene Parsing Network (MPSPNet) and UNet models applied to the fine automatic extraction of cropland in Shandong Province, and compared with the traditional object-oriented method, exploring the applicability of two deep convolutional neural network models in the automatic extraction of large-scale cropland. We also apply the trained models to the images of different regions and different time phases for the extraction of cropland, and explore the generalization ability of the models. The landscape features of cropland and uncertainty results are analyzed to explore the factors affecting the accuracy of cropland extraction by the models. Results show that: (1) MPSPNet and UNet models perform better than traditional object-oriented classification methods in the extraction of cropland at the district/county scale, the overall accuracy of the extraction of cropland at the provincial scale is better than 90% and there is no obvious difference between two models. (2) The landscape characteristic of cropland is an important factor that affects the effect of the two models, and the choice of the model has no obvious influence on the cropland extraction effect. The extraction effect is better in areas where the cropland landscape index is low and the plots are regular and flat, and the extraction effect is poor in the broken hilly areas of the plots with high cropland landscape index and in the noncropland plots whose characteristics are similar to the cropland, the UNet model is more likely to misclassify cropland in these areas. (3) The two models can obtain better cropland extraction effects in images of different regions and different time phases, and have strong generalization capabilities and temporal and spatial migration capabilities. This study proves the powerful feature learning capabilities of MPSPNet and UNET network models for high-resolution images, and the application potential of deep learning algorithms in fully automatic high-resolution cropland extraction.

Key words: Cropland, Remote Sensing, Convolutional neural network, MPSPNet, UNet

Supported by The National Key Research and Development Program of China (No. 2018YFC1504603) and the National High Resolution Earth Observation System (The Civil Part) Technology Projects of China (No. E03071112)