

基于卷积神经网络的高分遥感影像耕地提取研究

陈玲玲^{1,2} 施 政^{1,2} 廖凯涛^{1,2} 宋月君^{1,2} 张红梅³

(1. 江西省水利科学院, 南昌 330029; 2. 江西省土壤侵蚀与防治重点实验室, 南昌 330029;
3. 南昌工程学院水利与生态工程学院, 南昌 330099)

摘要: 高效精准地提取遥感影像中的耕地对农业资源监测以及可持续发展具有重要意义, 针对目前多数传统全卷积神经网络(FCN)模型在提取耕地时存在重精度而轻效率的缺陷, 本文建立基于 FCN 的轻量级耕地图斑提取模型(LWIBNet 模型), 并结合数学形态学算法进行后处理, 开展耕地图斑信息的自动化提取研究。该 LWIBNet 模型汲取了轻量级卷积神经网络和 U-Net 模型的优点, 以 Inv-Bottleneck 模块(由深度可分离卷积、压缩-激励块和反残差块组成)为核心, 采用高效的编码-解码结构为骨架, 将 LWIBNet 模型分别与传统模型的耕地提取效果、经典 FCN 模型的轻量性和精确度进行对比, 结果表明, LWIBNet 模型比表现最优的传统模型 Kappa 系数提高 12.0%, 比 U-Net 模型的参数量、计算量、训练耗时、分割耗时分别降低 96.5%、87.1%、78.2% 和 75%, 且 LWIBNet 的分割精度与经典 FCN 模型相似。

关键词: 耕地; 提取; 高分遥感影像; 卷积神经网络; LWIBNet

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2022)09-0168-10

OSID:



Cultivated Land Extraction from High Resolution Remote Sensing Image Based on Convolutional Neural Network

CHEN Lingling^{1,2} SHI Zheng^{1,2} LIAO Kaitao^{1,2} SONG Yuejun^{1,2} ZHANG Hongmei³

(1. Jiangxi Academy of Water Science and Engineering, Nanchang 330029, China

2. Jiangxi Key Laboratory of Soil Erosion and Prevention, Nanchang 330029, China

3. School of Hydraulic and Ecological Engineering, Nanchang Institute of Technology, Nanchang 330099, China)

Abstract: It is of great significance for agricultural resources monitoring to accurately extract cultivated land map information from remote sensing images. To improve the defects of traditional models for extracting cultivated land and solve the problem that most FCN model pays more attention to accuracy but ignores the consumption of time and computing resources, a lightweight model for extracting cultivated land map spots was established based on FCN (LWIBNet), and post-processing combined with mathematical morphology algorithm were used to carry out automatic extraction of cultivated land information. LWIBNet drew on the advantages of lightweight convolutional neural network and U-Net model, and it was built with the core of Inv-Bottleneck (composed of deep separable convolution, compression-excitation block and inverse residual block) and the skeleton of efficient coding-decoding structure. Compared LWIBNet with the cultivated land extraction effect of traditional model, and the computational resources and time consumption of classical FCN model. The results showed that LWIBNet was 12.0% higher than the Kappa coefficient of the best traditional model, and compared with U-Net, LWIBNet had 96.5%, 87.1%, 78.2% and 75% less parameters, calculation, training time and split time-consuming, respectively. Moreover, the segmentation accuracy of LWIBNet was similar to that of the classical FCN model.

Key words: cultivated land; extraction; high resolution remote sensing image; convolutional neural network; LWIBNet

收稿日期: 2022-03-16 修回日期: 2022-06-05

基金项目: 国家科技奖后备项目培育计划项目(20212AEI91011)和江西省水利科技项目(201922ZDKT08、202022YBKT20、202224ZDKT11、202123YBKT06)

作者简介: 陈玲玲(1995—),女,助理工程师,主要从事深度学习与遥感信息提取研究,E-mail: banjietaiyang@foxmail.com

通信作者: 宋月君(1982—),男,正高级工程师,博士,主要从事水土保持监测信息化研究,E-mail: well3292@126.com

0 引言

高分辨率遥感影像降低了中低分辨率影像的混合像元问题,能够较为清晰、准确地表达地物边界、形状、纹理信息、内部几何结构和空间关系^[1-2],当前高分遥感地物提取传统方法主要有最大似然法(Maximum likelihood, ML)^[3]、支持向量机法(Support vector machine, SVM)^[4-5]、随机森林法(Random forest, RF)^[6-7]、面向对象分类法^[8-9]等,并取得了较好的研究成果。然而传统方法大多是先由专业技术人员人工设定规则或划分样本,再根据自身算法或结合其他算法来实现分类提取^[10],受经验限制和人主观影响强,且高分影像中丰富的细节信息和复杂的规律往往无法被很好地描述,对特征差异大、信息复杂的遥感影像提取效果有限,容易导致椒盐噪声^[11]。随着遥感大数据时代的来临,特别是米级、亚米级高分遥感影像的出现,其拥有更清晰、更加丰富的地物影像信息^[12]。采用 SVM 方法将消耗大量的时间和计算资源,对计算机的性能要求高。虽然将大影像裁剪为若干小块来处理是一种解决办法,但这些提取方法需人工选取样本和设定参数,每一小块都要重新操作,繁琐耗时,还存在前后结果不一致的风险。目前,图斑精細化解译的实际过程依然以人工目视解译为主,遥感传统模型为辅的方式进行。

近年来,伴随着大数据、云计算和服务器的发 展,给深度学习(Deep learning, DL)技术带来了巨大的机遇^[13-15],卷积神经网络(Convolutional neural network, CNN)在遥感影像分类领域表现突出^[16],CNN 完全基于数据驱动,是一种无需专家知识来人工选择参数和设计特征,便可从数据中挖掘所需要的特征信息。具有智能化、自动化以及对海量遥感数据适应力强等特点。但 CNN 一般主要用作图像级的分类,归一化后输出结果为整个输入图像的一个数值描述(概率分布式),只能标识整幅图像类别,不能对标识图像中的每个像素点进行识别^[17]。因此,Long 等^[18]提出了全卷积神经网络(Fully convolution network, FCN),首次实现了端到端对图像进行像素级分类的功能,做到了像素对像素的映射。FCN 是目前许多语义分割方法的基础,相关学者基于 FCN 语义分割模型也开展耕地提取方面的研究工作^[19-21],基于 FCN 的模型算法,一旦训练出模型参数,即可实现复用推广,不必多次人工选定参数,大大节省了人工判读时间,实现了耕地的自动化提取,且最大限度地降低了人主观因素影响。然而,目前应用于遥感领域的 FCN 模型大多注重精度而

忽略了资源投入以及时间成本,在硬件资源和计算能力有限的情形下,可能会导致模型训练困难,耗时长,甚至无法实现等问题。

为解决上述问题,在以低资源消耗和高计算效率的轻量级卷积网络的基础上,通过对比 SqueezeNet^[22]、Xception^[23]、ShuffleNet^[24-25] 和 MobileNet^[26-27] 等优秀轻量级 CNN 模型,引入参数少、计算量小、精度高的深度可分离卷积、压缩-激励块和反残差块,并将这些块组合生成一个新模块(Inv-Bottleneck)。并借鉴语义分割模型实现端到端的高效框架^[28],引入 U-Net 网络结构的编码-解码框架^[29]。最后,以 Inv-Bottleneck 模块为核心,以编码-解码骨架为基础框架联合建立基于 FCN 的轻量级耕地图斑提取模型(Lightweight Inv-Bottleneck net, LWIBNet),应用数学形态学算法作为后处理手段,研究基于深度学习的高分遥感影像耕地图斑提取技术,以期土地利用变化以及生态环境效益评价提供数据和技术支撑。

1 研究方法

本文采用 LWIBNet 模型进行耕地图斑提取, LWIBNet 的整体架构见图 1。编码-解码结构是 LWIBNet 的基本框架,其中编码部分由 11 个 Inv-Bottleneck、2 个标准卷积和 1 个最大池化组成,并分为 5 个模块,将向解码部分输出 5 个不同尺度信息的特征图谱(FP),编码部分用于学习输入耕地图像的上下文语义信息并进行特征提取;解码部分由 4 个 Inv-Bottleneck、4 个标准卷积和 5 个上采样层组成,同样分为 5 个模块,将依次融合编码部分输出的 5 个 FP,形成更为丰富的特征信息。解码部分用于加强编码部分的特征提取结果,并恢复特征图的空间信息和分辨率,图 1 中, E-FP、D-FP 分别表示编码、解码部分的特征图谱,各 5 层。

解码部分的输出紧接带 Softmax 分类器的卷积层,此层能归纳出输入图像中具有相同语义的像素点,并输出一幅和输入图像大小相同的已分类标记好的图像,形成耕地图斑信息图。

1.1 Inv-Bottleneck 模块

Inv-Bottleneck 模块是本文模型的核心模块,由反残差块、压缩-激励块、深度可分离卷积和标准卷积构成,其能高效利用深层特征,在确保网络拥有一定准确率的同时降低参数数量和计算量,节约计算资源并提高效率,其结构图如图 2a 所示。NL 表示激活函数,即 ReLU6 和 H_swish。

1.1.1 深度可分离卷积

深度可分离卷积是轻量级神经网络模型的核心

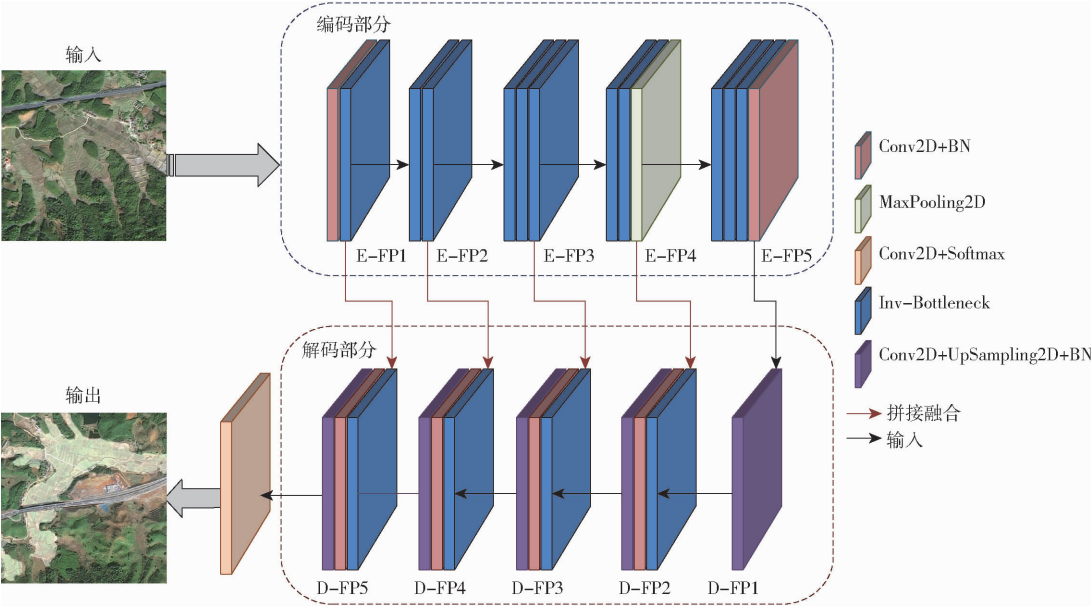


图1 LWIBNet 轻量级耕地地图斑提取模型结构图

Fig. 1 Structure diagram of LWIBNet lightweight cultivated land spot extraction model

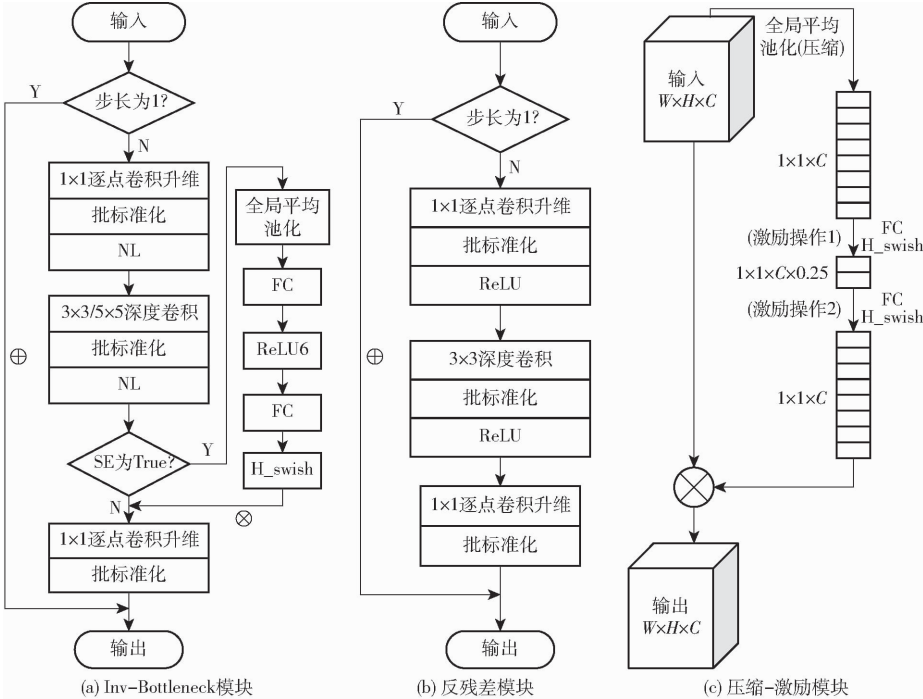


图2 模型组件结构图

Fig. 2 Structure diagrams of model components

工具,可显著降低计算量和计算成本,同时实现与标准卷积类似(或稍好)的性能,其将一个标准卷积拆分为两个独立的操作:深度卷积(DW)操作和逐点卷积(PW)操作^[30]。

深度卷积中,一个卷积核负责一个通道,一个通道只被一个卷积核所卷积,且只改变FP的大小,不改变通道数,相同卷积核的深度卷积比标准卷积需要更少的参数;逐点卷积等价于核大小为1×1的标准卷积,操作时只改变通道数,不改变FP,它常被用来关联特征图各通道之间的信息,进行信息交流。

使用DW和PW参数数量和计算量将显著减少,深度可分离卷积与标准卷积参数数量、计算量比值分别为

$$\frac{D_k D_k M + MN}{D_k D_k MN} = \frac{1}{N} + \frac{1}{D_k^2} \quad (1)$$

$$\frac{(D_k D_k M + MN) W_o H_o}{D_k D_k MN W_o H_o} = \frac{1}{N} + \frac{1}{D_k^2} \quad (2)$$

式中 D_k ——卷积核尺寸
 $W_o、H_o$ ——特征图宽度和高度
 M ——输入通道数 N ——输出通道数

若卷积核尺寸采用 3×3 , 深度可分离卷积的参数数量和计算量将减小为标准卷积的 $1/9 \sim 1/8$ 。在 Inv - Bottleneck 中混合使用 5×5 和 3×3 的 DW, 能使参数量不过多增加的同时有更好的效果, 并且为了降低过拟合的风险, 每步 DW 和 PW 之后都将使用批标准化 (BN)。

1.1.2 反残差模块

反残差模块是以 ResNet^[31] 的残差块为基础, 并结合深度可分离卷积优化而来的, 与 ResNet 残差块中维度先缩减后扩增相反, 它先使用 1×1 PW 将输入的低维 FP 升维, 然后在中间层利用轻量级的 DW 来提取特征并引入非线性, 再使用一个线性的 1×1 PW 将其映射回低维空间中, 如果步长为 1, 残差传播会被激活, 此时特征图将与输入图谱相加, 不仅能加深网络层数, 缓解训练时网络深度增加而出现的梯度消失问题, 还能减少参数量和计算量, 提高模型的效率和性能。具体结构见图 2b。

本文 Inv - Bottleneck 的骨架采用反残差结构, 并引入了 H_swish 激活函数^[32]。计算式为

$$H_swish(x) = x \frac{ReLU(x + 3)}{6} \quad (3)$$

因 H_swish 在更深层才能体现它的优势, 在编码部

分 Inv - Bottleneck 模块将在较后层才应用 H_swish 函数, 其他则继续使用 ReLU6。

1.1.3 压缩-激励模块

压缩-激励模块 (SE) 是能够让网络模型对特征进行校准的机制, 通过学习来自动获取每个特征通道的重要程度, 然后依照这一结果去提升有用的特征并抑制对当前任务用处小的特征^[33]。SE 能提升模型的质量, 但也会增加网络的总参数量和总计算量, 因此引入了改进的 SE 块, 其结构见图 2c, 包含 1 个压缩 (Squeeze) 和 2 个激励 (Excitation) 操作, 与未改进前相比, 第 1 个激励中增加了 0.25 的缩放参数, 可减少通道个数从而降低计算量, 两个激励中的激活函数都替换为计算量少的 H_swish。图中的 FC 表示全连接层, 可以看作 1×1 的卷积层, 有利于实现通道间的融合; W 、 H 和 C 分别表示特征图的宽、高和通道数。改进的 SE 模块被运用在 Inv - Bottleneck 模块的最后一层, 先进行 SE 操作再进行 PW 操作。实现压缩-激励块能在不破坏特征的情况下, 提升整个模型的运算性能。

1.2 编-解码结构

LWIBNet 的编码、解码部分具体结构和参数如表 1、2 所示。

表 1 编码部分具体结构和参数
Tab.1 Specific structure and parameters of encoding part

特征图谱	输入尺寸	操作运算	压缩-激励机制	激活函数	步长	通道数	输出尺寸
E - FP1	$256 \times 256 \times 3$	Conv + BN, 3×3	—	ReLU6	2	—	$128 \times 128 \times 16$
	$128 \times 128 \times 16$	Inv - Bottleneck, 3×3	—	ReLU6	1	16	$128 \times 128 \times 16$
E - FP2	$128 \times 128 \times 16$	Inv - Bottleneck, 3×3	—	ReLU6	2	72	$64 \times 64 \times 24$
	$64 \times 64 \times 24$	Inv - Bottleneck, 3×3	—	ReLU6	1	88	$64 \times 64 \times 24$
E - FP3	$64 \times 64 \times 24$	Inv - Bottleneck, 5×5	✓	H_swish	2	96	$32 \times 32 \times 40$
	$32 \times 32 \times 40$	Inv - Bottleneck, 5×5	✓	H_swish	1	240	$32 \times 32 \times 40$
	$32 \times 32 \times 40$	Inv - Bottleneck, 5×5	✓	H_swish	1	240	$32 \times 32 \times 40$
	$32 \times 32 \times 40$	Inv - Bottleneck, 5×5	✓	H_swish	1	120	$32 \times 32 \times 48$
E - FP4	$32 \times 32 \times 48$	Inv - Bottleneck, 5×5	✓	H_swish	1	144	$32 \times 32 \times 48$
	$32 \times 32 \times 48$	MaxPooling, 2×2	—	—	—	—	$16 \times 16 \times 48$
	$16 \times 16 \times 48$	Inv - Bottleneck, 5×5	✓	H_swish	1	288	$8 \times 8 \times 96$
E - FP5	$8 \times 8 \times 96$	Inv - Bottleneck, 5×5	✓	H_swish	1	576	$8 \times 8 \times 96$
	$8 \times 8 \times 96$	Inv - Bottleneck, 5×5	✓	H_swish	1	576	$8 \times 8 \times 96$
	$8 \times 8 \times 96$	Conv + BN, 1×1	—	H_swish	2	—	$8 \times 8 \times 96$

2 实验与结果分析

2.1 实验数据

本文收集了江西省奉新县、上高县、新余市渝水区、靖安县、崇仁县和宜丰县等 6 个县级行政区的高分 1 号以及高分 6 号遥感影像, 并经过大气校正、几何校正以及多波段影像融合处理后的分辨率为 2 m 的栅格影像以及经人工目视解译的耕地矢量数据作

为实验数据, 其中奉新县、上高县、新余市渝水区、靖安县和崇仁县的部分影像被截取作为训练影像, 相应的耕地矢量数据作为标签数据, 训练影像以及对应标签共同构成训练数据。为满足 GPU 显存需求和提升训练效率, 利用滑窗裁剪和随机裁剪, 将训练数据裁剪为若干 $256 \text{ 像素} \times 256 \text{ 像素}$ 的子图, 并按比例 4:1 分为训练集和验证集。另外, 为避免训练集影像过少引起的欠拟合或过拟合等问题, 还对训

表 2 解码部分具体结构和参数
Tab.2 Specific structure and parameters of decoding part

特征图谱	输入尺寸	操作运算	压缩-激励机制	激活函数	步长	通道数	输出尺寸
D - FP1	8 × 8 × 96	Conv, 2 × 2 + UpSampling, 2 × 2	—	ReLU	1	—	16 × 16 × 48
	16 × 16 × 48	Inv - Bottleneck, 3 × 3	✓	H_swish	1	144	16 × 16 × 48
D - FP2	16 × 16 × 48	Conv + BN, 3 × 3	—	ReLU	1	—	16 × 16 × 48
	16 × 16 × 48	Conv, 2 × 2 + UpSampling, 2 × 2	—	ReLU	1	—	32 × 32 × 40
D - FP3	32 × 32 × 40	Inv - Bottleneck, 3 × 3	✓	H_swish	1	240	32 × 32 × 40
	32 × 32 × 40	Conv + BN, 3 × 3	—	ReLU	1	—	32 × 32 × 40
	32 × 32 × 40	Conv, 2 × 2 + UpSampling, 2 × 2	—	ReLU	1	—	64 × 64 × 24
D - FP4	64 × 64 × 24	Inv - Bottleneck, 3 × 3	—	ReLU6	1	88	64 × 64 × 24
	64 × 64 × 24	Conv + BN, 3 × 3	—	ReLU	1	—	64 × 64 × 24
	64 × 64 × 24	Conv, 2 × 2 + UpSampling, 2 × 2	—	ReLU	1	—	128 × 128 × 16
D - FP5	128 × 128 × 16	Inv - Bottleneck, 3 × 3	—	ReLU6	1	16	128 × 128 × 16
	128 × 128 × 16	Conv + BN, 3 × 3	—	ReLU	1	—	128 × 128 × 16
	128 × 128 × 16	Conv, 3 × 3 + UpSampling, 2 × 2	—	ReLU	1	—	256 × 256 × 2

练集和验证集进行了水平翻转、垂直翻转、对角翻转等数据增广操作,并将其打乱使得样本分布更合理。最后生成了 77 969 幅 256 像素 × 256 像素图像的训练集,19 492 幅 256 像素 × 256 像素图像的验证集。宜丰县的遥感影像作为测试影像,测试 LWIBNet 模型提取耕地图斑的精确度。高分遥感影像相关信息如表 3 所示。

表 3 高分遥感影像的相关信息

Tab.3 Related information of high-resolution remote sensing images

区域	影像时相	影像分辨率/m
奉新县	1 月 (17%)、2 月 (3%)、3 月 (80%)	2
上高县	1 月 (25%)、2 月 (44%)、3 月 (31%)	2
新余市渝水区	1 月 (10%)、2 月 (90%)	2
靖安县	2 月 (5%)、3 月 (95%)	2
崇仁县	4 月 (100%)	2
宜丰县	1 月 (35%)、2 月 (65%)	2

2.2 实验方案与设置

选取 RF、SVM 和 ML 与 LWIBNet 模型的提取结果进行对比分析。其中,SVM 的每个类最大示例数设为 500;RF 的树设置为最大数量 100,最大深度 30;LWIBNet 模型的训练参数设置为:损失函数均采用交叉熵函数,优化器采用更优的 Nadam 优化算法^[34],分类器采用 Softmax 激活函数,初始学习率均采用 1×10^{-4} ,批尺寸为 36,验证集的损失率连续 10 轮没有下降,则停止训练。LWIBNet 与传统模型对比实验软硬件环境为:CPU 为 Intel (R) Xeon (R) CPU E5 - 2687W;GPU 为 NVIDIA GeForce GTX 1070;操作系统为 Windows 10;开发工具为 Tensorflow2.3 + Python3.8。

选取基于 FCN 的 U - Net 模型、Deeplabv3 + 模型(主干网络为 Resnet50)与 LWIBNet 的模型对轻量性和提取精确度进行了对比,实验环境为:CPU 为 Intel (R) Core i7 - 10875H;GPU 为 NVIDIA GeForce RTX 2070;操作系统为 Windows 10;开发工具为 Tensorflow2.3 + Python3.8。

为了避免传统模型以及基于 FCN 的模型产生细小图斑的影响,引入数学形态学方法,对传统分类模型和 LWIBNet 模型的提取结果进行后处理,传统模型先通过聚类处理将临近的类似分类区域聚类并进行合并,再用聚类之后的图像做过滤处理滤除小图斑。测试时增强算法具有提升精度的效果,所以 LWIBNet 模型的后处理为先使用增强算法来增强提取结果,再利用数学形态学的闭运算来优化图斑内部空洞,最后利用阈值法滤除杂波。

2.3 实验结果

使用优化后的耕地提取模型得到江西省宜丰县耕地图斑空间信息,如图 3 所示。

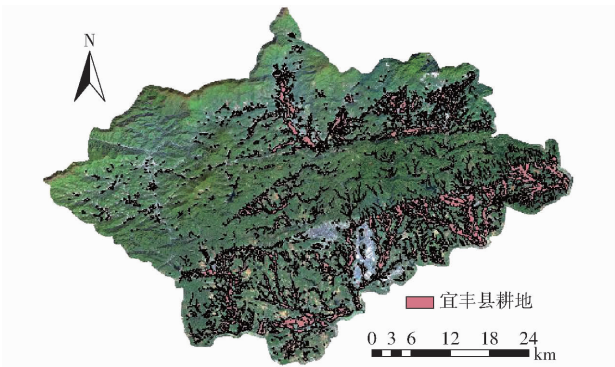


图 3 LWIBNet 模型提取的宜丰县耕地图斑空间分布信息图

Fig.3 Spatial distribution information map of cultivated land spots in Yifeng County extracted by LWIBNet

2.4 精度评价与分析

为验证 LWIBNet 模型提取耕地图斑的精度,选取宜丰县一块 7 107 像素 × 5 052 像素的区域作为测试影像,并将测试结果与相应区域人工标注的耕地信息图进行定性和定量比较。

2.4.1 传统模型与 LWIBNet 模型定性对比

图 4 为各模型从测试影像提取耕地图斑的整体对比,有助于全局性对各模型提取结果进行对比。图 5 为各模型从测试影像提取耕地图斑的细节对比,能更具体地辨别出正确分类和错误分类的地物。

由图 4 可知,从整体来看,除了 ML 将较多耕地错误识别为非耕地而导致整体分类情况和标签图相差较大外,其他模型的提取结果都与标签图在整体趋势上有较高吻合度。但与 LWIBNet 相比,没有经过后处理的传统分类模型都受到了严重的椒盐噪声干扰,从图 4d、4f、4h 可以看到后处理后的传统模型缓解了噪声对影像质量的影响,各图斑之间的界限也更明晰,但它们依然存在诸多错分漏分情况。而 LWIBNet 仅受少量杂波影响,其在未后处理的制图整洁度和美观度远优于经后处理的传统模型。在各模型中的后处理主要针对被错误识别为耕地的像素,后处理的效果在 LWIBNet 模型中的表现最为突出,错识别像素明显减少。根据图中错识别和漏识别像素数量和分布情况来看,ML 的表现最差,正确识别的像元数最少,RF 和 SVM 不管是后处理前还是后处理后分类表现相差不大,且这 3 个模型提取的图

斑都存在破碎度大和像元混合的问题,将本是成片的耕地识别成了破碎不连续的零散图斑,破坏了耕地信息的空间完整性。而 LWIBNet 模型结果不仅错分像素最少,提取的成片耕地大部分都具有连续性,后处理后的制图成果更加整洁,图 4k 中也可以看出,LWIBNet 模型提取的耕地区域与实际耕地的分布位置具有较好的一致性,从耕地图斑提取整体效果来看,LWIBNet 模型表现最优。

由图 5 可知,从细节分析上看,传统模型提取的图斑破碎度和像元混合看的更清晰,出现这些问题可能是因为耕地的光谱信息类内差异大,这与种植方式、土壤属性和作物类型等的不同有关。而 LWIBNet 模型能充分利用上下文语义信息,可减轻这种类内差异大带来的影响,故提取的图斑连续性好。传统模型还易将枯水期的河湖库塘、植被覆盖较低的其他林地裸土地等错误识别为耕地,这可能是受高分影像“同物异谱”和“异物同谱”影响。而 LWIBNet 模型完全趋于数据驱动,能反复多次地学习图像的底层和高层特征,这个优点可减轻上述现象的影响,故 LWIBNet 模型错识别为耕地的像素更少。从细节图能明显地看出后处理后的结果比处理前更加整洁和准确,但后处理也会带来一定的负面影响,容易滤除一些正确识别的孤岛小图斑,从而造成漏识别,但总体效果较好。通过与标签图像对比,ML 表现最差,SVM 略优于 RF,而 LWIBNet 表现最优。

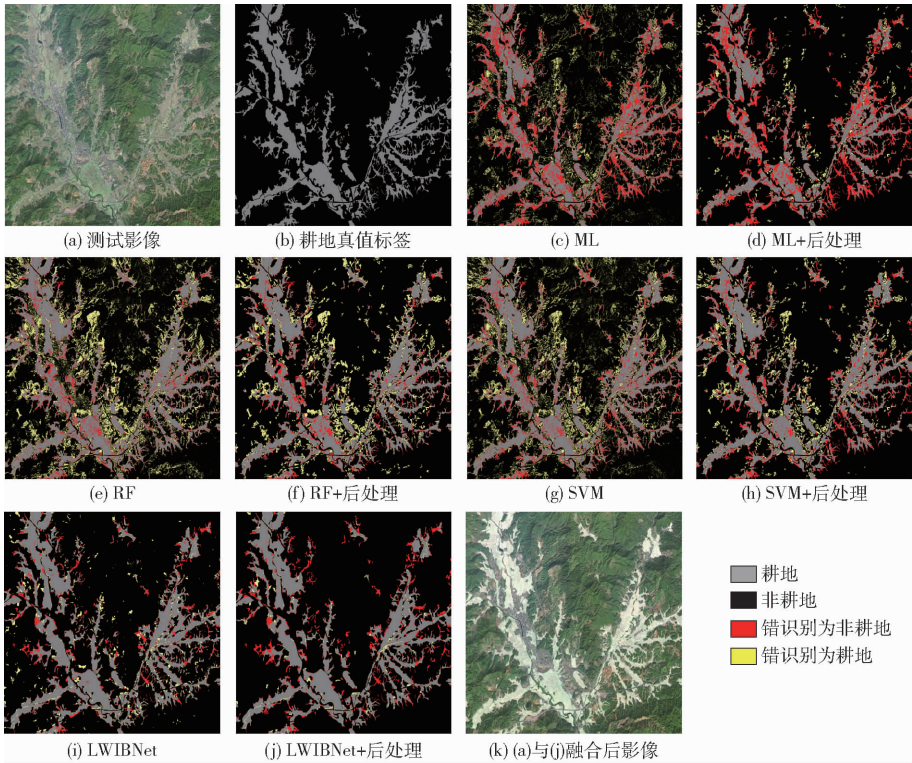


图 4 各模型从测试影像提取耕地图斑的整体对比

Fig.4 Overall comparisons of cultivated land patches extracted from test images by various models

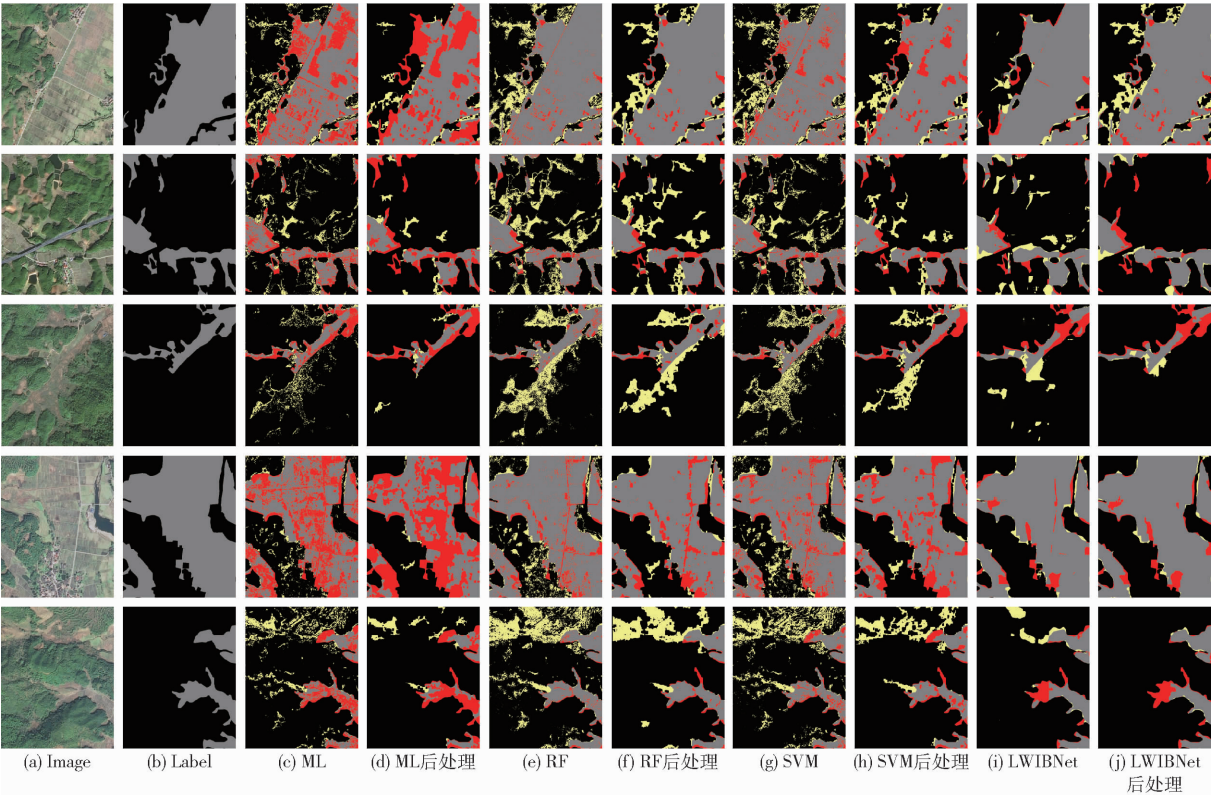


图5 各模型从测试影像提取耕地图斑的细节对比

Fig.5 Comparisons of details of cultivated land patches extracted from test images by different models

2.4.2 传统模型与 LWIBNet 模型定量对比

表4为LWIBNet模型与传统模型的定量评价结果。从表4可看出,所有模型后处理后的各项指标都高于处理前,ML、RF、SVM的Kappa系数比处理前分别提升了7.7%、9.6%和10.3%,而LWIBNet模型后处理后的提升较少,也提升了0.008。在3个传统模型中ML的评分最低,而SVM略高于RF。LWIBNet模型的各项评分显著优于传统模型,比表现最好的传统模型F1值提高11.6%,OA提高5.2%,MloU提高12.2%,Kappa系数提高12.0%。

表4 LWIBNet与传统模型的定量评价指标对比
Tab.4 Comparison of quantitative evaluation indexes between LWIBNet and traditional model

模型	F1 值	整体精度 (OA)	均交并比 (MloU)	Kappa 系数
ML	0.638	0.830	0.643	0.610
ML 后处理	0.681	0.878	0.688	0.657
RF	0.702	0.840	0.672	0.706
RF 后处理	0.759	0.883	0.735	0.774
SVM	0.701	0.847	0.676	0.705
SVM 后处理	0.764	0.892	0.743	0.778
LWIBNet	0.845	0.931	0.823	0.863
LWIBNet 后处理	0.853	0.938	0.834	0.871

通过对比 LWIBNet 提取的宜丰县耕地总面积

与人工解译的总面积发现,LWIBNet 提取的宜丰县耕地图斑面积为 278.30 hm²,人工解译的宜丰县耕地矢量数据中耕地面积为 300.59 hm²。绝对误差为 22.29 hm²,相对误差为 7.42%,模型提取的耕地图斑大体上和人工识别的一致,但一些林间耕地,窄小耕地等样本量少或复杂的耕地类型,容易被漏提取,但总体上表现较好,表明利用 LWIBNet 模型提取耕地图斑有较大的可行性。

2.4.3 经典 FCN 模型与 LWIBNet 模型的对比

实验主要包括模型轻量性和模型提取精确度对比。轻量性实验中包括模型的参数量、浮点运算数和训练耗时,这3个指标分别用于衡量模型复杂度、模型计算复杂度以及模型运行速度(表5)。对于模型提取精确度,主要开展了定性和定量的分析。此对比实验训练集和验证集是从大数据集中随机抽取7500幅图像,每个模型都训练40个周期(由于计算资源和时间限制,这里未进行充分训练,能对比本文模型与经典模型的轻量度、精确度的差异即可)。从表5可看出,在参数量上,LWIBNet为1.07 MB,U-Net的参数量是它的29倍,Deeplabv3+约为LWIBNet的11倍。在浮点运算数上,LWIBNet为14.14 GB,U-Net是LWIBNet的7.7倍,Deeplabv3+约为LWIBNet的3.7倍。在训练耗时上,U-Net的耗时最多,是Deeplabv3+的2.3倍,是LWIBNet的4.6倍,LWIBNet的耗时最少,每步耗时139 ms。

表 5 LWIBNet 与经典 FCN 模型的计算资源、时间消耗指标对比(轻量度对比)

Tab.5 Comparison of computing resources and time consumption between LWIBNet and classical FCN model

模型	参数量/ MB	浮点运算 数/GB	训练耗时/ ms	分割耗时/ ms
Deeplabv3 +	11. 85	52. 57	277	9
U - Net	31. 06	109. 30	637	20
LWIBNet	1. 07	14. 14	139	5

在分割耗时上,U - Net 的每步耗时最多,LWIBNet 的耗时最少。LWIBNet 无论是参数量、计算量还是训练耗时都低于经典语义分割模型,符合轻量级网络模型的特点。

图 6 为各深度学习模型从测试影像提取耕地图斑的整体对比。由图 6 可知,Deeplabv3 + 的红色像素比黄色像素更多,表明错误识别的像素中错误识别为

非耕地的更多,U - Net 的红色像素比黄色像素略多一点,两种错误识别概率相近,LWIBNet 主要是黄色像素比红色像素多,也就是错误识别为耕地的更多一点;3 个模型对比来看,LWIBNet 模型的黄色像素最多,也错误识别为耕地的更多,但是它红色像素最少,正确提取的灰色像素也比其他模型多一点,Deeplabv3 + 红色像素最多,黄色最少,U - Net 的黄色和红色像素数量居于中间。定量对比如表 6 所示。从表 6 可看出,U - Net 与 LWIBNet 在各项指标上只有 0.001 ~ 0.005 的差异,LWIBNet 与 Deeplabv3 + 也仅有 0.005 ~ 0.028 的差异,且 LWIBNet 的大部分指标更优于 Deeplabv3 +。综合定性和定量结果,LWIBNet 与经典 FCN 模型有着相似的分割精度,且该模型具有一定轻量性,它的计算资源和时间消耗更低。

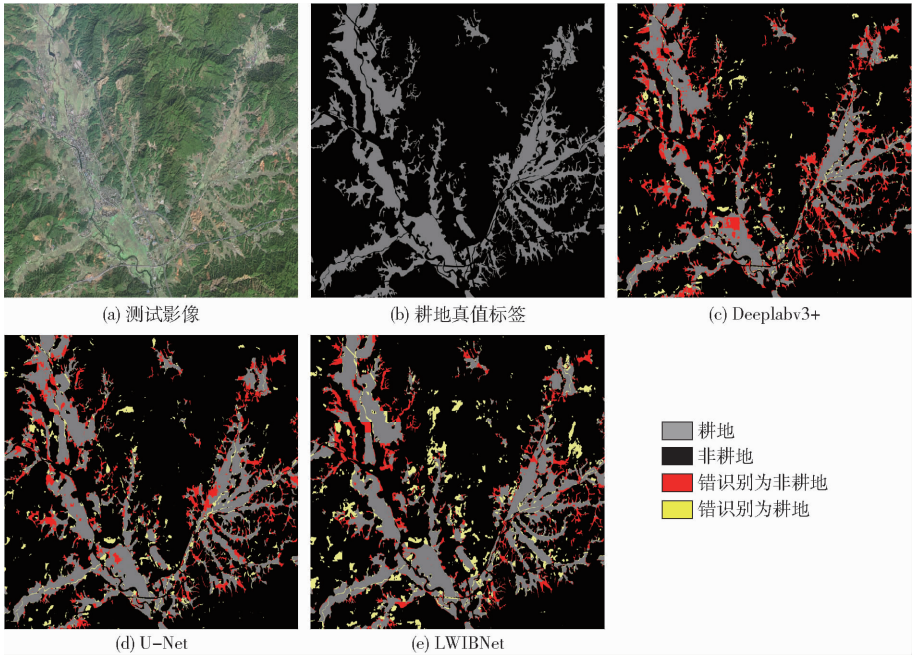


图 6 各深度学习模型从测试影像提取耕地图斑的整体对比

Fig.6 Each deep learning model extracted whole contrast of cultivated land patches from test images

表 6 LWIBNet 与经典 FCN 模型的提取结果定量对比

Tab.6 Quantitative comparison of extraction results between LWIBNet and classical FCN model

模型	F1 值	整体精度 (OA)	均交并比 (MIoU)	Kappa 系数
U - Net	0. 757	0. 896	0. 742	0. 768
Deeplabv3 +	0. 738	0. 896	0. 732	0. 741
LWIBNet	0. 756	0. 891	0. 738	0. 769

3 结论

(1)综合轻量级卷积神经网络和 U - Net 的优点,以 Inv - Bottleneck 模块(由深度可分离卷积、压缩-激励块和反残差块组成)为核心,联合高效的编

码-解码结构为框架,建立了基于 FCN 的轻量级耕地图斑提取模型(LWIBNet 模型)。

(2)LWIBNet 模型有效改善了传统模型提取结果破碎的问题,并减轻了椒盐噪声影响,其 F1 值、整体精度(OA)、均交并比(MIoU)和 Kappa 系数等各项指标均比表现最好的传统模型(经后处理的 SVM)分别提高了 11.6%、5.2%、12.2% 和 12.0%;且 LWIBNet 模型训练得到模型参数具有更好的普适性,降低了人为主观因素的干扰,自动化水平更高。

(3)在相同的实验环境和设置下,LWIBNet 模型的参数量、计算量和训练耗时都低于经典的 FCN 模型 Deeplabv3 + 和 U - Net,其中,Deeplabv3 + 的参

数量是LWIBNet模型的11倍,U-Net的计算量是LWIBNet模型的7.7倍,U-Net的训练耗时是LWIBNet模型的4.6倍,LWIBNet无论是参数量、计算量还是训练耗时和分割耗时都低于经典的FCN

模型Deeplabv3+和U-Net,且提取结果的精度与经典FCN模型相似,表明LWIBNet能以更少的计算资源和时间消耗实现与经典FCN模型相似的分割精度。

参考文献

- [1] 熊曦柳,胡月明,文宁,等.耕地遥感识别研究进展与展望[J].农业资源与环境学报,2020,37(6):856-865.
XIONG Xiliu, HU Yueming, WEN Ning, et al. Research progress and prospect of cultivated land recognition by remote sensing [J]. Journal of Agricultural Resources and Environment, 2020,37(6):856-865. (in Chinese)
- [2] 刘大伟,韩玲,韩晓勇.基于深度学习的高分辨率遥感影像分类研究[J].光学学报,2016,36(4):306-314.
LIU Dawei, HAN Ling, HAN Xiaoyong. Research on classification of high-resolution remote sensing images based on deep learning[J]. Acta Optica Sinica, 2016,36(4):306-314. (in Chinese)
- [3] 刘焕军,杨昊轩,徐梦园,等.基于裸土期多时相遥感影像特征及最大似然法的土壤分类[J].农业工程学报,2018,34(14):132-139,304.
LIU Huanjun, YANG Haoxuan, XU Mengyuan, et al. Soil classification based on multi-temporal remote sensing image characteristics and maximum likelihood method in bare soil period[J]. Transactions of the CSAE, 2018,34(14):132-139, 304. (in Chinese)
- [4] KADAVI P R, LEE C. Land cover classification analysis of volcanic island in aleutianarcusing anartificial neural network (ANN) and a support vector machine (SVM) fromlandsat imagery[J]. Geosciences Journal, 2018, 22(4): 652-665.
- [5] PHAN T N, KAPPAS M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery[J]. Sensors, 2018, 18(1): 1-20.
- [6] 于新洋,赵庚星,常春艳,等.随机森林遥感信息提取研究进展及应用展望[J].遥感信息,2019,34(2):11-17.
YU Xinyang, ZHAO Gengxing, CHANG Chunyan, et al. Research progress and application prospect of remote sensing information extraction of random forest[J]. Remote Sensing Information, 2019, 34(2):11-17. (in Chinese)
- [7] ZHANG X M, HE G J, ZHANG Z M, et al. Spectral-spatial multi-feature classification of remote sensing big data based on a random forest classifier for land cover mapping[J]. Cluster Computing, 2017, 20(3): 2311-2321.
- [8] 范莹琳,娄德波,张长青,等.基于面向对象的铁尾矿信息提取技术研究——以迁西地区北京二号遥感影像为例[J].自然资源遥感,2021,33(4):153-161.
FAN Yinglin, LOU Debo, ZHANG Changqing, et al. Research on information extraction technology of iron tailings based on object-oriented——taking Beijing No. 2 remote sensing image of Qianxi area as an example[J]. Remote Sensing of Natural Resources, 2021,33(4):153-161. (in Chinese)
- [9] LIU T, ABD-ELRAHMAN A. Multi-view object-based classification of wetland land covers using unmanned aircraft system images[J]. Remote Sensing of Environment, 2018, 216(6): 122-138.
- [10] 董蕴雅,张倩.基于CNN的高分遥感影像深度语义特征提取研究综述[J].遥感技术与应用,2019,34(1):1-11.
DONG Yunya, ZHANG Qian. Summary of research on depth semantic feature extraction of high-scoring remote sensing images based on CNN[J]. Remote Sensing Technology and Application, 2019,34(1):1-11. (in Chinese)
- [11] 李越帅,郑宏伟,罗格平,等.集成U-Net方法的无人机影像胡杨树冠提取和计数[J].遥感技术与应用,2019,34(5):939-949.
LI Yueshuai, ZHENG Hongwei, LUO Geping, et al. Extraction and counting of *Populus euphratica* crown from UAV images using U-Net method[J]. Remote Sensing Technology and Application, 2019,34(5):939-949. (in Chinese)
- [12] 孙家波,张晓艳,牛鲁燕,等.基于高分辨率遥感影像的耕地信息快速提取方法研究[J].山东农业科学,2018,50(3):132-136,141.
SUN Jiabo, ZHANG Xiaoyan, NIU Luyan, et al. Study on fast extraction method of cultivated land information based on high-resolution remote sensing images[J]. Shandong Agricultural Sciences, 2018,50(3):132-136,141. (in Chinese)
- [13] DONG S, WANG P, ABBAS K. A survey on deep learning and its applications[J]. Computer Science Review, 2021, 40 (100379):1-22.
- [14] 薛金林,李雨晴,曹梓建.基于深度学习的模糊农田图像中障碍物检测技术[J].农业机械学报,2022,53(3):234-242.
XUE Jinlin, LI Yuqing, CAO Zijian. Obstacle detection based on deep learning for blurred farmland images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(3):234-242. (in Chinese)
- [15] 张宏鸣,张国良,朱珊珊,等.基于U-Net的葡萄种植区遥感识别方法[J].农业机械学报,2022,53(4):173-182.
ZHANG Hongming, ZHANG Guoliang, ZHU Shanna, et al. Remote sensing recognition method of grape planting regions based on U-Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(4):173-182. (in Chinese)
- [16] 许慧敏.基于深度学习U-Net模型的高分辨率遥感影像分类方法研究[D].成都:西南交通大学,2018.
XU Huimin. Research on classification method of high-resolution remote sensing images based on deep learning U-Net model [D]. Chengdu: Southwest Jiaotong University, 2018. (in Chinese)

- [17] 张祥甫,刘健,石章松,等.基于深度学习的语义分割问题研究综述[J].激光与光电子学进展,2019,56(15):20–34.
ZHANG Xiangfu, LIU Jian, SHI Zhangsong, et al. A survey of semantic segmentation based on deep learning[J]. Laser & Optoelectronics Progress, 2019, 56(15):20–34. (in Chinese)
- [18] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431–3440.
- [19] 刘巍,吴志峰,骆剑承,等.深度学习支持下的丘陵山区耕地高分辨率遥感信息分区分层提取方法[J].测绘学报,2021,50(1):105–116.
LIU Wei, WU Zhifeng, LUO Jiancheng, et al. Hierarchical extraction method of high-resolution remote sensing information of cultivated land in hilly and mountainous areas supported by deep learning[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(1):105–116. (in Chinese)
- [20] KHAN A H, FRAZ M M, SHAHZAD M. Deep learning based land cover and crop type classification: a comparative study[C]//2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), 2021:1–6.
- [21] 宁纪锋,倪静,何宜家,等.基于卷积注意力的无人机多光谱遥感影像地膜农田识别[J].农业机械学报,2021,52(9):213–220.
NING Jifeng, NI Jing, HE Yijia, et al. Recognition of film-covered farmland by UAV multispectral remote sensing images based on convolution attention[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9):213–220. (in Chinese)
- [22] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[J]. arXiv preprint arXiv:1602.07360, 2016:1–13.
- [23] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1251–1258.
- [24] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[J]. arXiv:1707.01083, 2017:1–9.
- [25] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//European Conference on Computer Vision, 2018:122–138.
- [26] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv:1704.04861, 2017:1–9.
- [27] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 4510–4520.
- [28] 张哲晗,方薇,杜丽丽,等.基于编码-解码卷积神经网络的遥感图像语义分割[J].光学学报,2020,40(3):46–55.
ZHANG Zhehan, FANG Wei, DU Lili, et al. Semantic segmentation of remote sensing image based on encoding-decoding convolutional neural network[J]. Acta Optica Sinica, 2020, 40(3):46–55. (in Chinese)
- [29] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015: 234–241.
- [30] KAISER L, GOMEZ A N, CHOLLET F. Depthwise separable convolutions for neural machine translation[J]. arXiv preprint arXiv:1706.03059, 2017:1–10.
- [31] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- [32] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 1314–1324.
- [33] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011–2023.
- [34] DOZAT T. Incorporating nesterov momentum into adam[C]//4th International Conference on Learning Representations, 2016: 2013–2016.