



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法
作者: 任诗曼, 朱军, 方铮, 李闯农, 梁策, 谢亚坤, 李维炼, 张天奕
DOI: 10.13203/j.whugis20210251
收稿日期: 2021-12-19
网络首发日期: 2022-01-04
引用格式: 任诗曼, 朱军, 方铮, 李闯农, 梁策, 谢亚坤, 李维炼, 张天奕. 联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法[J/OL]. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20210251>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

引用格式：

任诗曼, 朱军, 方铮, 等. 联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法[J]. 武汉大学学报-信息科学版, 2021, DOI: 10.13203/j.whugis20210251 (REN Shiman, ZHU Jun, FANG Zheng, et al. Woodland Extraction of SPOT7 Image Based on Multi-scale Attention Mechanism and Edge Constraint[J]. Geomatics and Information Science of Wuhan University, 2021, DOI: 10.13203/j.whugis20210251)

联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法

任诗曼¹ 朱军¹ 方铮² 李闯农¹ 梁策³ 谢亚坤¹ 李维炼¹ 张天奕¹

¹ 西南交通大学地球科学与环境工程学院, 四川 成都, 611756

² 四川省国土科学技术研究院, 四川 成都, 610045

³ 中国铁道科学研究院集团有限公司 电子计算技术研究所, 北京, 100081

摘要：林地是国家重要的自然资源和经济资源，掌握林地分布状况对林地资源调查管理具有重要意义。针对传统林地提取方法精度较低且边界不规则的问题，本文设计一种联合多尺度注意力机制与边缘约束的林地提取方法，首先，构建一种端到端的多尺度注意力神经网络模型，充分提取影像中林地的上下文特征，对不同尺度下的林地进行语义描述，实现高精度的林地像素级表达；其次，构建边缘约束规则，对提取结果进行边界优化，提高林地提取结果的可读性。为证明方法的有效性，以四川省绵阳市三台县作为试验区，建立数据集并进行林地提取试验，结果显示，本文方法提取的精确度为 81.9%，召回率为 75.6%，F1 为 78.1%，交并比为 64.2%。结果证明，本文方法在遥感影像林地提取应用上具有更好的效果。

关键词：林地提取；多尺度；注意力机制；边缘约束

林地是国家重要的自然资源和战略资源，严格保护林地、发展森林资源是保障国土生态安全的基本要求，而掌握林地分布是了解林地现状、制定环境政策、保护林地资源的数据基础^[1]。随着遥感技术的迅速发展，遥感影像为林地信息调查提供了理想数据，如何利用遥感影像实现林地信息的准确提取具有重要意义。

现有方法为遥感影像林地提取提供了不同的思路，主要可分为两类：一类是传统的林地提取方法，一类是基于深度学习的林地提取方法^[2-4]。传统方法的思路是将影像中单个像素或其临近像素作为处理单元，依据影像光谱和纹理特征提取林地，例如使用归一化植被指数、比值植被指数

获取林地光谱特征，再通过分类器进行林地信息提取^[5-6]，但这类方法着眼于局部像素而忽略邻近整片地块的纹理和结构信息，造成提取精度低。由于林地上树种多样，在遥感影像中表现出不同的光谱特征，存在明显的同物异谱现象，使像素级林地提取方法的精度低且椒盐效应明显^[7]。在此之后，研究人员通过面向对象的方法来解决这一问题，该方法利用地物光谱特征和形状特征提取同质区域，并对各个区域进行特征分析以实现地物信息提取^[8-9]。但高分辨率遥感影像特征极为复杂，在复杂背景下的遥感影像中存在单株立木、耕地等噪声，使得面向对象的方法提取结果中林地斑块细碎；同时，此类方法是基于光谱

收稿时间：2021-12-19

项目资助：四川省自然资源厅项目(KJ-2020-4)；国家自然科学基金(U2034202, 41871289)；四川省科技计划资助(2020JDTD0003)。

第一作者：任诗曼，硕士生，主要从事遥感图像信息提取研究。Email: rens9604@163.com

通讯作者：朱军，博士，教授，主要从事三维地理系统与虚拟地理环境研究。Email: vgezj@163.com

特征的像素聚合，在影像分割过程中无法充分考虑林地的上下文特征。在进一步的研究中，由于机器学习方法对人工提供的特征依赖少，可以通过特征提取函数，学习样本完成参数拟合，逐渐被用于遥感影像林地提取，例如基于支持向量机的遥感影像林地分类、基于决策树的橡胶林地识别等^[10-11]。然而，由于影像中树种类型多样、相互混淆，机器学习方法于深层次语义特征提取方面仍有欠缺，而通过构建深层次神经网络则可以通过自动学习的方法有效抽取影像深层次特征，可以进一步解决以上问题。

近年来，基于深度学习的遥感影像分类算法已被验证在分类准确率上优于传统算法^[12-14]，尤其在像素级语义分割方面，全卷积神经网络具有较高的优势，亦可以用于遥感影像分割任务，如使用半监督全卷积网络进行城市土地覆盖分类与变化检测^[15]；基于无人机图像数据，使用端到端的卷积神经网络提取森林树种信息^[16]；结合 POI-Net(Position-of-Interest Network)和 DeepUNet(Deep U-connection Network)网络检测林地，提升对不同尺度林地的感知能力^[17]。尽管以上研究在深度学习应用于遥感领域的进展明显，然而，相对于一般地物影像中的林地更加特殊，首先，林地上树种多样，在遥感影像中表现出不同的光谱特征，同物异谱特征明显；其次，林地光谱特征与耕地、草地等相似，同谱异物现象严重；最后，由于林地中包括大面积森林覆盖和林盘中小面积林地，在遥感影像中的多尺度特征更加明显；此外，现阶段缺乏林地图像分割数据集，使深度学习在该领域的研究受到限制。

针对以上问题，本文提出一种联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法，首先，在编码阶段联合空间与通道注意力建立上下文信息抽取模块，揭示遥感影像中林地与其他地物上下文语义特征关联关系，提高网络对于遥感影像中地物异质性特征的感知能力，在解码阶段利用金字塔结构，增强模型对不同种类

林地和不同尺度林地的特征提取能力；其次，为进一步优化林地边界规则化表达，解决像素级语义分割生成的多边形锯齿现象严重的问题，构建边缘约束规则，获取与林地真实边界更为契合的轮廓结果；最后，开展案例实验分析，利用 SPOT7 影像制作数据集，进行林地提取实验，并将本文方法与多种神经网络模型对比，以证明本文方法的有效性。

1 林地提取方法

考虑到林地在遥感影像中呈现出尺度差异、同谱异物特征、同物异谱特征，本文针对性地提出联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法，如图1所示，首先，引入通道和空间注意力模块，生成不同通道和不同区域的重要性权重，减弱对背景区域和非显著性特征的关注；其次，设计多尺度注意力机制的林地提取网络，获取不同尺度特征并进行融合操作，使特征图对林地的语义信息和细节信息都有更强的表示能力；最后，建立边缘约束规则，对林地提取结果进行后处理，填补图斑细小孔洞并优化提取结果中林地地块的边缘，平滑林地边界。

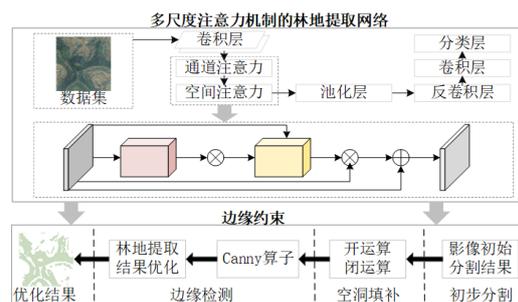


图1 总体框架

Fig.1 The Overall Framework

1.1 多尺度注意力网络架构

多尺度注意力林地提取网络框架如图2所示。该网络是一种端到端的对称结构，首先，在编码阶段，以 ResNet(Residual Network)^[18]作为网络特征提取的基线结构，

提取影像中深层次信息，同时在每一步池化操作前引入通道注意力和空间注意力，分别获取不同通道和空间位置像素的权重参数，通过对应位置相乘将学习的权重参数叠加到各个特征点位置，充分聚合局部信息和全局信息，捕捉林地与其他地物的上下文语义特征关联关系，增强网络特征提取能力和场景理解能力；其次，在解码

阶段，通过反卷积处理还原图像尺寸至输入影像大小，解决池化操作后特征图分辨率降低的问题，同时，将池化层与反卷积层进行串行跳层连接，结合低层局部特征和高层语义特征，使输出特征图具有更好的林地特征表示能力，并利用多尺度特征聚合结构，融合多种不同层次的特征信息，实现林地的多尺度特征提取。

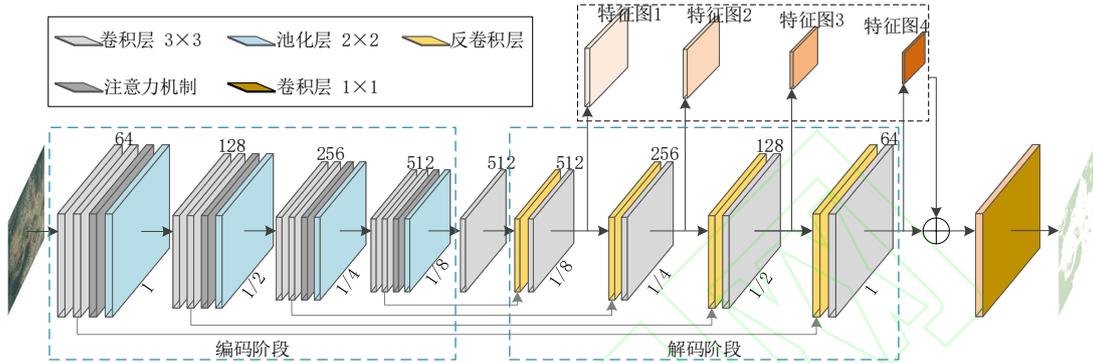


图2 多尺度注意力林地提取网络结构

Fig.2 The Structure of the Multi-scale Attention Woodland Extraction Network

1.1.1 融合上下文特征的注意力机制

遥感影像中的上下文特征是指多种地物对象之间位置和特征表达的关联关系，其能够直接或者间接影响地物信息提取结果^[19-20]。通过分析实验区遥感影像中林地与其他对象之间的关联关系，将其上下文特征分为三个方面。首先，林地特征表达与草本植物空间位置上具有一定的相邻性，

如林地常与耕地、草地等地物伴随出现；其次，在光谱特征表达方面，林地与耕地及草地具有较明显的光谱相似性；最后，林地内部及侧面具有较明显的阴影，草本植物则由于比较低矮因而阴影不明显，具有一定的语义差异性。本文通过融合双重注意力机制，提高网络对于林地上下文信息的抽取能力，有效降低林地特征提取过程的不确定性，如图3所示。

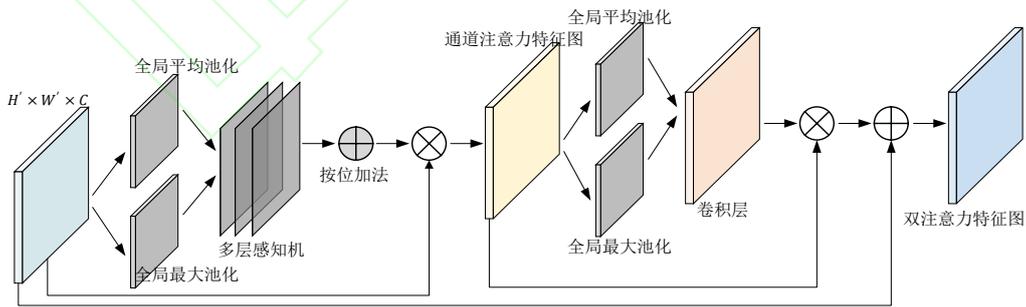


图3 双重注意力机制

Fig.3 Dual Attention Mechanism

在网络的编码阶段，在每一次最大池化操作前引入注意力模块，将大小为 $H' \times W' \times C$ 的特征图 $F_{1,2}$ 作为输入，分别经过全局平均池化和全局最大池化，可由公式 (1) 和公式 (2) 表示

$$Z_{\text{avg}}^C = F_{\text{avg}}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j) \quad (1)$$

$$Z_{\text{max}}^C = F_{\text{max}}(U_c) = \max_{1 \leq i \leq H} \left(\max_{1 \leq j \leq W} U_c(i, j) \right) \quad (2)$$

其中， Z_{avg}^C 、 Z_{max}^C 分别表示经过全局

平均池化和全局最大池化压缩后 C 个通道的全局信息, F_{avg} 、 F_{max} 表示特征权重提取操作, U_c 为输入特征图在第 c 通道的特征, i 和 j 为特征图在空间维的坐标。全局平均池化和全局最大池化聚合了特征映射的空间信息。为了减少网络的训练参数, 将池化后的特征经过一个共享多层感知机以压缩通道数, 然后使用按位加法融合输出特征向量, 可由公式 (3) 表示

$$F_c = \text{sigmoid}(W_2\psi(W_1Z_{avg}^c) + W_2\psi(W_1Z_{max}^c)) \quad (3)$$

其中, C 为输入层和输出层神经元个数, F_c 为通道注意力特征权重; ψ 为ReLU(Rectified Linear Unit)激活函数, W_1 、 W_2 为多层感知机的权重。为了增强网络的特征处理能力, 在局部特征表示上建立丰富的上下文关联关系, 加入全局相关性特征, 在通道注意力的基础上引入空间注意力, 由公式 (4) 表示

$$F_s = \text{sigmoid}(\phi^{7 \times 7}(Z_{avg}^s \text{cat} Z_{max}^s)) \quad (4)$$

其中, ϕ 表示卷积操作, cat 表示通道拼接操作, Z_{avg}^s 、 Z_{max}^s 分别为经过空间注意力模块的全局平均池化和全局最大池化压缩后的全局信息。输入通道注意力处理后的结果 F_c , 分别对 F_c 进行通道维度的全局平均池化和全局最大池化, 得到两个二维特征描述, 采用通道拼接的方法融合这两个特征描述, 然后经过一个卷积层, 使用 sigmoid 函数获取权重系数 F_s 。最后, 联合空间与通道注意力特征权重, 得到双重注意力机制特征图 F_{CS} , 如公式 (5) 所示

$$F_{CS} = F_s \cdot F_c + F_{l2} \quad (5)$$

特征图 F_{CS} 中每个特征值是经过双重注意力机制处理后加权的特征值与原始特征值相加所得, 因而 F_{CS} 包含了不同区域林地的上下文关联关系, 通过注意力机制捕捉林地上下文特征, 从而提高网络对于林地信息的提取能力。

1.1.2 林地多尺度特征聚合

遥感影像中林地尺度差异大, 例如大面积的苗圃和林盘中的小面积林地。神经

网络浅层由于下采样次数少, 特征感受野小, 适合处理较小的林地地块; 相对来说深层网络则具有更大的特征感受野, 适合检测覆盖面积较广的林地。通过多尺度特征融合结构, 如图4所示, 在不同感受野下提取林地特征, 将解码阶段不同抽象层级的特征进行融合, 有助于捕获不同尺度林地的局部和全局特征, 解决单一尺度特征表达不足的问题。编码阶段将影像分辨率依次下采样为输入影像的 1/2、1/4、1/8, 解码阶段使用反卷积逐级上采样, 恢复输入影像中的空间信息和林地边缘信息, 保留解码阶段上采样过程中不同尺度的特征图 1 至特征图 4, 然后使用四层金字塔模块, 将特征图 1、特征图 2、特征图 3 通过双线性插值法上采样至与特征图 4 相同的尺寸, 再使用 1×1 的卷积对每层结果进行降维, 改变特征图的通道数使各个尺度的处理结果通道数相同, 最后将特征图 4'、特征图 3'、特征图 2'、特征图 1'逐层相加相同位置元素进行通道融合, 通过聚合 4 个不同层次的林地特征, 可以提高网络对不同尺度林地的提取效果。

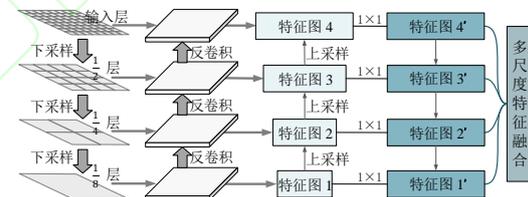


图4 多尺度特征聚合结构

Fig.4 The Structure of Multi-scale Feature Aggregation

1.2 边缘约束的林地分割结果优化

遥感影像中林地分布零散且林地和耕地交错, 而耕地在遥感影像上的光谱特征与林地相似, 这导致分割结果在林地边缘处易出现错提或漏提问题, 而且由于受到单株立木和林间裸地的影响使提取结果中存在细碎图斑和细小孔洞。针对林地分割结果中边界不清晰、图斑细碎等情况, 使用基于数学形态学的方法优化林地分割结果, 具体流程如图5所示, 结合形态学闭运算和开运算滤除林地提取结果的噪声, 填

补图斑中不规则空洞、消除细碎图斑，然后通过 Canny 检测算子提取分割结果中林地边界，产生连续闭合边缘，在闭合边缘约束下，对图像分割得到的林地提取结果进行融合分析，使用图斑面积作为区域合并准则，将连续面积较小的图斑（本文中设置为面积小于 20 平方米）与其有最长公共面积的相邻图斑合并，进一步消除细碎林地图斑以及林地图斑内被误判为非林地的细小孔洞，再创建指数多项式逼近图斑边界，产生连续、光滑的边界曲线，对边缘进行平滑处理。边缘优化后的林地分割结果可以反馈到样本数据集中，重新训练模型，提高样本的准确性。

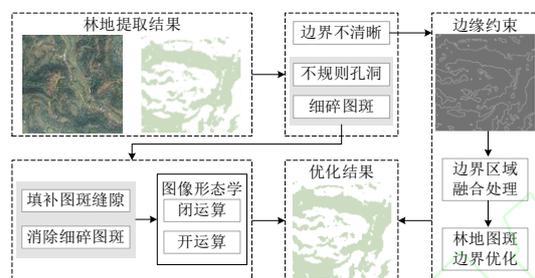


图 5 林地分割结果优化

Fig.5 Optimization of Woodland Segmentation Results

2 试验与分析

2.1 研究区概况

为了验证上述提出的方法，本文选择四川省绵阳市三台县作为研究区，采用 2017 年获取的空间分辨率 2m 的 SPOT7 影像作为试验数据，如图 6 所示。该县林地资源丰富，全县林业用地面积 8.96 万公顷，森林覆盖率 38.8%^[21]。该区域地势北高南低，林地地块大小不一、树种类型繁多（毛竹、银杏、柏树等），是我国西南林区的一个代表性区域。

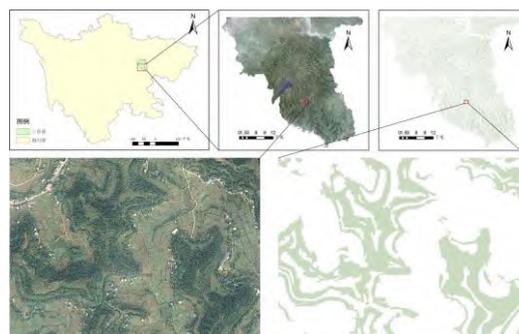


图 6 研究区概况

Fig.6 The Overview of Study Area

2.2 试验数据集

获取神经网络的参数需要大量的训练数据，而获取大量数据的成本高，往往采用数据增强的方式增加训练数据量。数据增强结果如图 7 所示，其中图(a)为未经处理的原始影像，图(b)至图(f)分别为经过旋转 90°、亮度增强、水平翻转、随机裁剪以及添加高斯噪声处理后的影像。本文所用数据集包含 10756 张训练影像、2752 张验证影像和 2019 张测试影像，大小均为 512×512 像素。

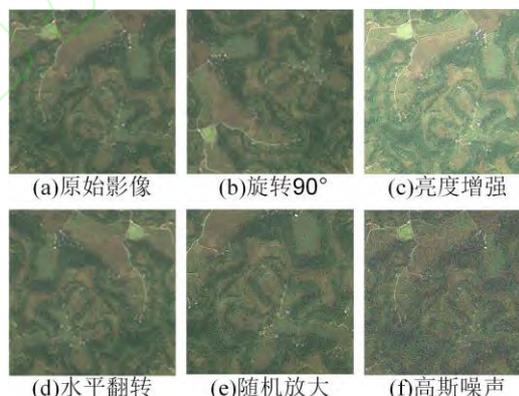


图 7 数据增强

Fig.7 Data Augmentation

2.3 试验环境

遥感影像林地提取试验平台为 Windows10 系统。深度学习软件框架为 TensorFlow1.13.1，Python 版本为 3.6。试验所用的硬件配置为：Intel(R) Core(TM) i7-9700K CPU @ 3.60GH 处理器，GeForce RTX 2070 SUPER 显卡，显存 8G，16G 内

存。

神经网络内部参数通过训练模型迭代获取，而训练时还需要人为设置部分超参数。经过反复测试，本文设置初始学习率为 0.001，通过监测 Val_loss 的值，当经过 5 个 epoch 而性能不提升时学习率降低为原来的 0.5，训练批次大小为 4，训练阶段迭代 100 次。

2.4 林地提取精度评价标准

遥感影像地物提取实际是图像语义分割的一种应用，所得的结果通常采用图像语义分割的评价方法，因此，本文使用正确率 (Accuracy, Acc)、交并比 (Intersection of Union, IoU)、F1、精确度 (Precision, P)、召回率 (Recall, R) 五种评价指标，计算公式分别为

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (7)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (8)$$

$$P = \frac{TP}{TP+FP} \quad (9)$$

$$R = \frac{TP}{TP+FN} \quad (10)$$

其中，TP(True Positive)表示正确提取的林地像元；FP(False Positive)表示错判为林地的非林地像元；TN(True Negative)表示正确提取的非林地像元；FN(True Negative)表示错判为非林地的林地像元。

2.5 试验结果与分析

为了验证本文方法的有效性，将本文方法与 FCN8s(VGG16)(Fully Convolutional Network)^[12]、UNet(U-connection Network)^[22]、RCF(ResNetXt-CBAM-FPN)^[23-25]进行对比，并对各个模型输出的结果进行定性分析和定量分析。其中，RCF 模型是在 ResNetXt-50(Aggregated Residual Network)^[23]的基础上，加入 CBAM(Convolutional Block Attention Module)^[24]模块和 FPN(Feature Pyramid Network)^[25]模块。

2.5.1 林地提取结果分析

1) 定性分析

为了评价本文方法的性能，对本文方法进行定性分析。FCN8s(VGG16)、UNet、RCF 和本文方法的林地提取结果如图 8 所示。从图 8(c)、8(d)可以得出，FCN8s(VGG16)、RCF 网络林地提取结果，边界较为模糊，林地图斑内存在较多孔洞，其中 FCN8s(VGG16)网络漏提林地明显，而 RCF 网络错提林地明显。从图 8(e)可以得出，UNet 网络比 FCN8s(VGG16)和 RCF 林地提取效果好，但其提取结果不够精细，容易将林地周围的耕地或草地错分类为林地。从图 8(f)可以得出，本文方法获取的结果，林地图斑更贴近图 8(b)标签，面积较小的林地也能较好识别。

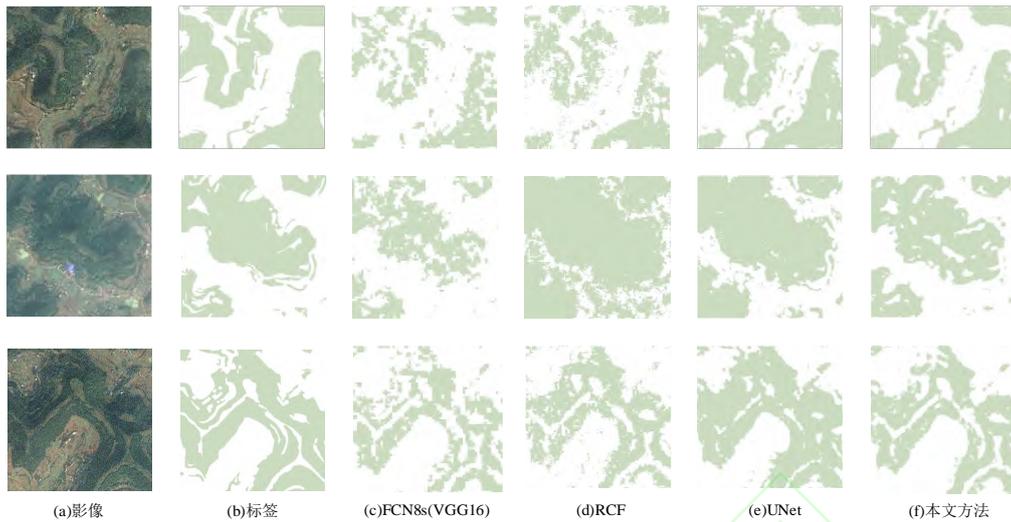


图8 不同方法的林地提取结果图像对比

Fig.8 Image Comparison of Different Methods for Woodland Extraction

2) 定量分析

为了定量评价本文林地提取方法，采用 FCN8s(VGG16)、RCF、UNet 的林地提取结果作为参考，利用 Accuracy、IoU、F1、Precision、Recall 五个指标评价遥感影像林地提取结果的精度，其精度如表1所示。由表1可知，本文方法在林地提取任务上与其他三种方法相比，本文方法的精度最高，

其林地提取的 F1 为 77.4%，Recall 为 74.2%，IoU 为 63.2%，与 FCN8s(VGG16)相比较分别高出了 15.9%、24.1%、16.9%，与 RCF 相比较分别高出了 15.3%、14.0%、17.3%，与 UNet 相比较分别高出了 3.2%、2.6%、1.9%。通过定量评价验证本文提出的多尺度注意力林地提取网络能够有效提高林地提取精度，优于其它三种方法。

表1 不同方法的林地提取结果精度对比

Tab.1 Accuracy Comparison of Different Methods for Woodland Extraction

| 方法 | Acc (%) | P (%) | R (%) | IoU (%) | F1 (%) |
|--------------|---------|-------|-------|---------|--------|
| FCN8s(VGG16) | 76.1 | 81.2 | 50.1 | 46.3 | 61.5 |
| RCF | 71.7 | 73.1 | 60.2 | 45.9 | 62.1 |
| UNet | 80.3 | 78.3 | 71.6 | 61.3 | 74.2 |
| 本文方法 | 82.1 | 82.0 | 74.2 | 63.2 | 77.4 |

2.5.2 优化后林地提取结果分析

1) 定性分析

本文所用数据集林地分布零散且林地与耕地交错，而耕地在遥感影像上的光谱特征与林地相似，这导致分割结果在林地边缘处易出现边缘表达不清晰，图斑边缘

呈锯齿状，如图9(c)所示。采用边缘约束的林地分割结果优化方法进行处理，其结果如图9(d)所示。对比图9(c)和图9(d)可知，图9(d)中红色矩形框内的林地图斑无细小空洞且边界较平滑，即优化处理能够消除提取结果中的细碎图斑并平滑图斑边界，使地块更完整、边界更清晰。

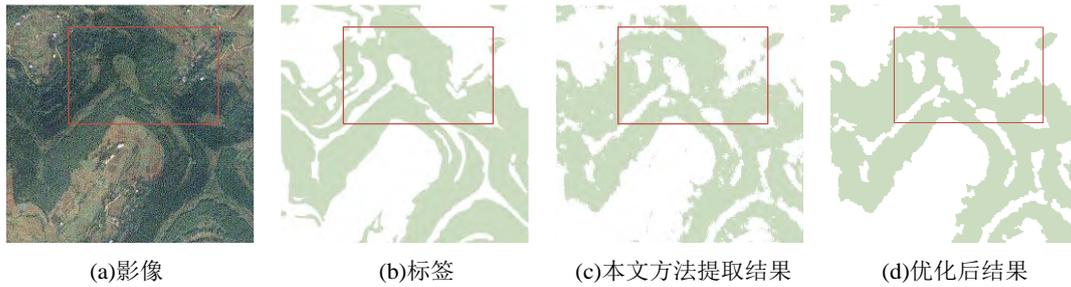


图9 林地提取结果优化前后图像对比

Fig.9 Image Comparison of Woodland Extraction Results before and after Optimization

2) 定量分析

将未优化的林地提取结果与优化后的林地提取结果进行多指标精度对比, 定量分析边缘约束的林地分割结果优化方法的有效性。林地提取结果优化处理前后精度对比如表2所示, 从表2中可以得出, 优化

处理后的林地提取结果 F1、Recall 和 IoU 分别提高至 78.1%、75.6%和 64.2%, 与优化前的林地提取结果相比较, 分别提高了 0.7%、1.4%和 1.0%。可见对林地提取结果进行优化处理能够小幅度提升遥感影像林地提取的精度。

表2 林地提取结果优化前后精度对比

Tab.2 Accuracy Comparison of Woodland Extraction Results before and after Optimization

| 方法 | Acc (%) | P (%) | R (%) | IoU (%) | F1 (%) |
|-------|---------|-------|-------|---------|--------|
| 本文方法 | 82.1 | 82.0 | 74.2 | 63.2 | 77.4 |
| 优化后结果 | 82.5 | 81.9 | 75.6 | 64.2 | 78.1 |

3 结语

1) 针对遥感影像林地尺度差异大、林地上树种多样导致的同物异谱问题以及像素级语义分割生成的林地图斑边缘锯齿问题, 针对性地提出了一种联合多尺度注意力机制与边缘约束的遥感影像林地提取方法, 为林地精准提取提供了技术支撑。

2) 在覆盖四川省绵阳市三台县地区的 SPOT7 遥感影像数据集上的试验结果表明, 使用本文方法提取林地的召回率为 75.6%, 交并比为 64.2%, F1 为 78.1%, 正确率为 82.5%, 精确度为 81.9%。

本文提出的林地提取方法通过双重注意力机制充分提取影像中林地上下文语义特征, 提高了网络的深层次特征表达, 有效解决影像中林地同物异谱及同谱异物问题, 并联合多尺度特征聚合, 提高网络对不同尺度林地的感知能力, 在此基础上对提取结果进行边缘优化处理, 解决林地图斑边缘锯齿的问题并有效填补图斑中细小孔洞, 对精准掌握林地分布状况具有重要应用价值。

参考文献

- [1] General Office of the State Council of the People's Republic of China. The executive meeting of the State Council passed the outline of the national plan for the protection and utilization of Woodland[Z]. http://www.gov.cn/ldhd/2010-06/09/content_1623921.htm, 2010 (国务院办公厅. 国务院常务会议通过全国林地保护利用规划纲要[Z]. http://www.gov.cn/ldhd/2010-06/09/content_1623921.htm, 2010)
- [2] Chen Zhou, Fei Xianyun, Gao Xiangwei, et al. Extraction of Urban Green Space with High Resolution Remote Sensing Image Segmentation[J] *Bulletin of Surveying and Mapping*, 2020(12): 17-20 (陈周, 费鲜芸, 高祥伟, 等. 高分辨率遥感影像分割的城市绿地提取研究[J]. 测绘通报, 2020(12): 17-20)
- [3] Xu Qing. Research on Vegetation Extraction of High Resolution Remote Sensing Image Based on Attention Model[D]. *Wuhan University*, 2020 (徐青. 基于注意力机制的高分辨率遥感影像植被提取研究[D]. 武汉大学, 2020)
- [4] Huang Jie. Determining Various Forest Types from

- GaoFen-2 Satellite Image Using Object based CNN[D]. *China University of Geosciences (Beijing)*, 2020 (黄杰. 基于 OBIA-CNN 的高分二号卫星影像林地类型识别[D]. 中国地质大学(北京), 2020)
- [5] Sun Jianguo, Ai Tinghua, Wang Pei, et al. Assessing Vegetation Degradation Based on NDVI-climate Variables Feature Space[J]. *Geomatics and Information Science of Wuhan University*, 2008, 33(6): 573-576 (孙建国, 艾廷华, 王沛, 等. 基于 NDVI-气候变量特征空间的植被退化评价[J]. 武汉大学学报·信息科学版, 2008, 33(6): 573-576)
- [6] Cheng K, Wang J. Forest Type Classification Based on Integrated Spectral-spatial-temporal Features and Random Forest Algorithm—A Case Study in the Qinling Mountains[J]. *Forests*, 2019, 10(7)
- [7] Kim C, Hong S. The Characterization of a Forest Cover through Shape and Texture Parameters from QuickBird Imagery[C]. *IEEE International Geoscience and Remote Sensing Symposium*, 2008, 3: 692-695
- [8] Wang Chunyan, Liu Jiabin, Xu Aigong, et al. A New Method of Fuzzy Supervised Classification of High Resolution Remote Sensing Image[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(6): 922-929 (王春艳, 刘佳新, 徐爱功, 等. 一种新的高分辨率遥感影像模糊监督分类方法[J]. 武汉大学学报·信息科学版, 2018, 43(6): 922-929)
- [9] Cheng Shiyao, Mei Tianchan, Liu Guoying. Application of Multi-level MRF Using Structural Feature to Remote Sensing Image Classification[J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(9): 1180-1187 (程诗尧, 梅天灿, 刘国英. 顾及结构特征的多层次马尔科夫随机场模型在影像分类中的应用[J]. 武汉大学学报·信息科学版, 2015, 40(9): 1180-1187)
- [10] Hou Yichen, Zhao Pengxiang, Yang Zhiwei, et al. Forest Classification and Accuracy Assessment in ZY3 Image with SVM Method[J]. *Journal of Northwest Forestry University*, 2016, 31(1): 180-185 (侯逸晨, 赵鹏祥, 杨伟志, 等. 基于 SVM 的资源三号影像林地分类及精度评价研究[J]. 西北林学院学报, 2016, 31(1): 180-185)
- [11] Liu Xiaona, Feng Zhiming, Jiang Luguang, et al. Application of Decision Tree Classification to Rubber Plantations Extraction with Remote Sensing[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2013, 29(24): 163-172 (刘晓娜, 封志明, 姜鲁光. 基于决策树分类的橡胶林地遥感识别[J]. 农业工程学报, 2013, 29(24): 163-172)
- [12] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4): 640-651
- [13] Kattenborn T, Leitloff J, Schiefer F, et al. Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 173: 24-49
- [14] Huang B, Zhao B, Song Y. Urban Land-use Mapping Using a Deep Convolutional Neural Network with High Spatial Resolution Multispectral Remote Sensing Imagery[J]. *Remote Sensing of Environment*, 2018, 214: 73-86
- [15] Ji Shunping, Tian Siqi, Zhang Chi. Urban Land Cover Classification and Change Detection Using Fully Atrous Convolutional Neural Network[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(2): 233-241 (季顺平, 田思琦, 张驰. 利用全空洞卷积神经网络进行城市土地覆盖分类与变化检测[J]. 武汉大学学报·信息科学版, 2020, 45(2): 233-241)
- [16] Schiefer F, Kattenborn T, Frick A, et al. Mapping Forest Tree Species in High Resolution UAV-Based RGB-imagery by Means of Convolutional Neural Networks[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 170: 205-215
- [17] Gui Y, Li W, Wang Y, et al. Woodland Detection Using Most-sure Strategy to Fuse Segmentation

- Results of Deep Learning[C]. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019: 1-4
- [18] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 770-778
- [19] Gong Yiping. Context Aware CNN for Object Detection from VHR Remote Sensing Imagery[D]. *Wuhan University*, 2018 (宫一平. 基于上下文感知的高分辨率遥感影像目标检测[D]. 武汉大学, 2018)
- [20] Ri Changyong. Research on Semantic Image Classification Based on Context Information[D]. *Zhejiang University*, 2014 (李昌英. 基于上下文信息的语义图像分类研究[D]. 浙江大学, 2014)
- [21] Li Jian. Development Status and Countermeasures of Forestry Industry in Santai County[J]. *Rural Science and Technology*, 2018(16): 50-51 (李建. 三台县林业产业发展现状及对策[J]. 乡村科技, 2018(16): 50-51)
- [22] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, 9351: 234-241
- [23] Xie S, Girshick R, Dollár P, et al. Aggregated Residual Transformations for Deep Neural Networks[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 5987-5995
- [24] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module[C]. *European Conference on Computer Vision (ECCV)*, 2018: 3-19
- [25] Lin T, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 936-944

Woodland Extraction of SPOT7 Image Based on Multi-scale Attention Mechanism and Edge Constraint

REN Shiman¹ ZHU Jun¹ FANG Zheng² LI Chuangnong¹ LIANG Ce³ XIE Yakun¹

LI Weilian¹ ZHANG Tianyi¹

¹ Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China

² Sichuan Institute of land Science and Technology, Chengdu 610045, China

³ Institute of Electronic Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China

Abstract: As woodland is an important natural and economic resource of China, it is important to understand the distribution of woodland for the investigation and management of woodland resources. In this paper, we design a woodland extraction method combining multi-scale attention mechanism and edge constraint to tackle the issue of low accuracy and irregular boundaries in traditional forest extraction methods. First, an end-to-end multi-scale attentional neural network model is constructed to fully extract the context features of woodland in remote sensing images, and semantically describe woodland at different scales to achieve high-precision pixel-level expression of woodland. Secondly, the edge constraint rules are constructed to optimize the boundary of the

extraction results, to improve the readability of the extraction results. To prove the effectiveness of the proposed method, Santai County, Mianyang City, Sichuan Province, is taken as the experimental area to establish datasets and carry out woodland extraction experiments. The results show that the extraction accuracy of this method is 81.9%, the recall rate is 75.6%, F1 is 78.1%, IoU(Intersection of Union) is 64.2%, and the method in the paper has a better effect in the application of woodland extraction with remote sensing image.

Key words: woodland extraction; multiscale; attention mechanism; edge constraint

First author: REN Shiman, Master, specializes in remote sensing image information extraction. E-mail: rens9604@163.com

Corresponding author: ZHU Jun, PhD, professor, specializes in three-dimension Geographic Information System and virtual geographical environment. E-mail: vgezj@163.com

Foundation support: Project of Department of Natural Resources of Sichuan Province(KJ-2020-4); the National Natural Science Foundation of China(U2034202, 41871289) ; Sichuan Science and Technology Program(2020JDTD0003).

网络首发:

标题: 联合多尺度注意力机制与边缘约束的SPOT7影像林地提取方法

作者: 任诗曼, 朱军, 方铮, 李闯农, 梁策, 谢亚坤, 李维炼, 张天奕

收稿日期: 2021-12-19

DOI: 10.13203/j.whugis20210251

引用格式:

任诗曼, 朱军, 方铮, 等. 联合多尺度注意力机制与边缘约束的 SPOT7 影像林地提取方法[J]. 武汉大学学报·信息科学版, 2021, DOI: 10.13203/j.whugis20210251 (REN Shiman, ZHU Jun, FANG Zheng, et al. Woodland Extraction of SPOT7 Image Based on Multi-scale Attention Mechanism and Edge Constraint[J]. Geomatics and Information Science of Wuhan University, 2021, DOI: 10.13203/j.whugis20210251)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

利用具有注意力的 Mask R-CNN 检测震害建筑物立面损毁

眭海刚, 黄立洪, 刘超贤

武汉大学学报·信息科学版, 2020, 45(11): 1660-1668

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20200158>

应急遥感制图中敏感目标自动检测与隐藏方法

鲁鹏杰, 许大璐, 任福, 徐胜华, 邱天奇, 彭瑞

武汉大学学报·信息科学版, 2020, 45(8): 1263-1272

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20200131>