

doi: 10.6046/gtzyyg.2020278
引用格式: 武宇,张俊,李屹旭,等. 基于改进 U-Net 的建筑物集群识别研究[J]. 国土资源遥感,2021,33(2):48-54. (Wu Y, Zhang J, Li Y X, et al. Research on building cluster identification based on improved U-Net[J]. Remote Sensing for Land and Resources, 2021, 33(2):48-54.)

基于改进 U-Net 的建筑物集群识别研究

武 宇¹, 张 俊¹, 李屹旭², 黄康钰¹

(1. 贵州大学矿业学院, 贵阳 550025; 2. 贵州大学农学院, 贵阳 550025)

摘要: 针对 U-Net 在高分影像建筑物提取中部分建筑物边缘特征易模糊或丢失的问题, 提出一种对高分影像建筑物边缘增强, 同时对 U-Net 部分卷积过程进行改进的优化的建筑物提取方法。首先利用域变化递归滤波的方式对建筑物边缘进行增强, 将增强后影像输入 U-Net 神经网络中进行训练; 其次为充分利用建筑物在高分影像上丰富的细节特征, 尝试在原 U-Net 结构基础上, 从训练图像和标签中提取成对的补丁以增加训练数据, 这些补丁进一步加强了正反向深度学习中建筑物高维特征的获取; 最后在影像上实现建筑物提取。对辽宁省盘锦市邻接渤海湾地区 2017 年 9 月 29 日高分二号影像建筑物提取实验结果表明, 对于包含阴影区域干扰较多的非理想样本数据, 用 U-Net 识别建筑物得到的整体分类精度为 75.99%, 而改进方法最高整体分类精度可达 83.12%, 较原 U-Net 网络精度提高 7.13 百分点, 证明该方法行之有效。

关键词: 深度学习; 域变化递归滤波; U-Net; 边缘增强; 建筑物提取

中图法分类号: TP 751 **文献标志码:** A **文章编号:** 1001-070X(2021)02-0048-07

0 引言

现代科学技术的日益发展,城市建筑物外观设计也随之趋于复杂,表观信息丰富,致使传统低分辨率遥感影像已远远不能表达现代建筑物的精细特征。高空间分辨率卫星的升空使得利用高分影像对建筑物分割成为可能^[1-2]。然而,尽管高分辨率遥感影像丰富的细节特征在一定程度上增大了目标地物的类内差异、减小了类间差异,但是“同物异谱,同谱异物”的现象却变得更加普遍、明显,同类地物在影像上的形态与特征可能大相径庭,直接加大了类内混合像元及阴影等因素的影响。为此,一些学者曾引入一些模式识别算法,如最大似然分类法(maximum likelihood classification, MLC)和支持向量机(support vector machine, SVM),此后,一些学者提出了一种基于数学形态的合成核 SVM 算法^[3-5],利用空间统计因素作为数据源,通过不同的合成核构造方式将空间信息和光谱信息融合,引入遥感图像分

类中。但是,这些算法仅利用图像的光谱信息,忽略了其空间结构纹理的作用,因而在提高分类精度的过程中还存在局限性,常常会导致提取精度过低、边界不完整等问题。另外,图像中的噪声如传感器产生的热噪声、脉冲噪声和数字化过程中出现的量化噪声等,也会对图像压缩、边缘检测以及图像分割过程造成严重影响^[6]。

近年来,逐步兴起的深度学习算法在计算机视觉领域中带来的显著效益引起了大量研究者的关注。深度学习算法以数据作为驱动,更加适用于目前遥感影像数据大量积累的情况,且由于其在图像分割中的出色表现,很快被引入遥感影像地物提取中,并已取得了不错的效果^[7-8]。例如方旭等^[9]证明利用改进的全卷积神经网络(fully convolutional networks, FCN)可做出优良的结果,较 SVM 算法提升效果约 10%; 2012 年, Krizhevsky 等^[10]提出了名为 AlexNet 的卷积神经网络(convolutional neural networks, CNN)结构,并在 ImageNet 比赛中以领先第二名 10% 的成绩获得冠军; Mou 等^[11]首次将循环

收稿日期: 2020-09-04; 修订日期: 2021-01-11
基金项目: 贵州省科学技术基础研究计划项目“基于 GPS 的地壳弹性形变反演模型研究”(编号: 黔科[2017]1054)、国家自然科学基金项目“基于地表拓扑特征的无控制点矿山变形监测与预警”(编号: 41701464)和贵州大学研究生创新基地建设项目“测绘科学与技术研究生创新实践基地建设项目”(编号: 贵大研 CXJD[2014]002)共同资助。
第一作者: 武 宇(1995-),男,硕士研究生,目前主要从事 InSAR 地形变监测与提取、基于深度学习的遥感影像分割理论学习和研究。
Email: ywu@niglas.ac.cn。
通信作者: 李屹旭(1980-),女,硕士,讲师,主要从事测绘数据处理、遥感科学与技术及地理信息系统教学、科研工作。Email: lyxgis@163.com。

神经网络(recurrent neural network, RNN)应用在高光谱图像分类中,较 SVM 达到了更高的精度;陈睿敏等^[12]利用 U - Net 深度学习模型对红外遥感影像进行地物分类,较传统机器学习方法提升 6%。

各类深度学习模型中,U - Net 模型由于学习能力强、易于扩展、整体性能稳定、鲁棒性强等特点而备受青睐^[13-17]。但是,U - Net 网络模型最初被用于检测医学影像,而高分遥感影像不同于医学影像,不仅具有光谱特性,还具有非常复杂的空间结构、纹理特征和上下文关系,如果简单地采用 U - Net 模型,则很可能在最大池化采样中丢失一些细节信息,尤其是图像边缘特征的损失,从而导致分类失真^[18]。为此,本文提出首先采用域变换递归滤波(domain transform recursive filter, DTRF)对高分影像进行预处理,在一定程度上抑制噪声并尽量保证影像边缘信息,以此确保 U - Net 神经网络对高分影像更正确的分割;同时,尝试采用从训练图像和标签中提取成对的补丁以增加训练数据对 U - Net 编码 - 解码环节进行改进,这些补丁可望进一步加强正反向深

$$|t(x_i, I(x_i)) - t(x_j, I(x_j))| = \| (x_i, I(x_i)) - (x_j, I(x_j)) \| \quad (1)$$

式中: $| |$ 为绝对值; $\| \|$ 为范数; x_i, x_j 分别为均匀采样得到的任意两点。因此, t 只需要保留最近邻上的两点在 C 上的距离,然后选取信号中某一区间 ct , 令 $ct(x) = t(\hat{x}) = t(x, I(x))$, 那么所有的变换必须在 1 范数下满足:

$$ct(x + h) - ct(x) = h + |I(x + h) - I(x)| \quad (2)$$

其中得到的 $\mathbf{R}$ 中的相邻样本之间欧式距离必须等于 $\mathbf{R}^2$ 间的距离, h 在此处为信号上两相邻采样间隔,为简化计算,通常假设 ct 单调递增,上式两端同除以 h 且取极限趋向 0 得:

$$ct'(x) = 1 + |I'(x)| \quad (3)$$

那么在 Ω 中的 (u, w) 两点间的距离可表示为:

$$ct(w) - ct(u) = \int_u^w 1 + |I'(x)| dx \quad (4)$$

可见,曲线 ct 在 1 范数下的弧长被保留了下来,而所有点之间的测地距离可以通过式(4)计算。

在实际滤波过程中,为了控制滤波器的尺寸和模糊度,常常将域变换定义为应用近似距离变换,即对于给定的一维信号 I ,域变换被定义为:

$$U_i = I_0 + \sum_{i=1}^n (1 + \frac{\delta_i}{\delta_r} |I_i - I_{i-1}|) \quad (5)$$

式中: U_i 为域变换信号; δ_r 和 δ_s 分别为与值域标准差(控制滤波器模糊度)和与信号空域标准差(控制

度学习中建筑物高维特征的获取,以此提高建筑物识别精度,为方便简记该方法为 DTRF - Unet。

1 基于 DTRF - Unet 方法对高分遥感影像建筑物的提取原理

1.1 DTRF 滤波

DTRF 是一种实时的边缘保留滤波,其在改善图像分类的性能中非常有效,一维 DTRF 对信号经过平滑后可以很好地去噪声,而且主要边缘信息也可以得到保留^[19-21]。滤波过程可分为域变换和递归滤波两步,DTRF 过程如下:

假设在 $I: \Omega \rightarrow \mathbf{R}, \Omega = [0, + \infty] \subset \mathbf{R}$ 的作用域内存在某一维信号,在 $\mathbf{R}^2$ 定义域上存在某信号 $C: (x, I(x))$, 目的是找到一个变化 $t: \mathbf{R}^2 \rightarrow \mathbf{R}$, 在 $\mathbf{R}$ 中保留信号 C 上任意点之间的原始距离,之后在 Ω 上均匀采样得到 $x \in \{x_0, x_1, x_2, \dots, x_n\}$, $ct(x) = t(\hat{x}) = t(x, I(x))$, 其中 $x_{i+1} = x_i + h$, h 为采样间隔。 t 满足:

滤波窗口尺寸); n 为像元个数,输入信号通过递归滤波器进行处理,即

$$J_i = (1 - a^b)I_i + a^b J_{i-1} \quad (6)$$

式中: J_i 为第 i 个像元的滤波输出; a 为反馈系数; b 为变换域中两个相邻样本的距离。当图像接近边缘时, b 不断增大, a^b 趋于 0,迭代运算终止,从而达到边缘保留的目的。

1.2 U - Net 神经网络

U - Net 神经网络模型由 Ronneberger 等于 2015 年首次提出^[22-23],最初应用在医学影像处理,因其形状恰似“U”形而得名。U - Net 是一种典型的编码 - 解码(encoder to decoder)结构,在 encoder 模块使用卷积(conv)与最大池化(maxpooling)操作提取图像的低级特征,在 decoder 模块使用向上解码(up - sampling)与 conv 恢复输出图像(feature map)的分辨率提取高维特征,在相对应的 conv 之间融合特征信息。up - sampling 仅仅是将抽象高维特征恢复到模板的过程,解码在将编码压缩数据恢复的时候,特征 scale 会发生变化,这时横向连接的 skip connection 起到了补充信息的作用,在高层补充了语义的信息,在底层则细化了分割的轮廓等^[24-25]。所以从特征中已经大致恢复了区域的分割信息,再加上原始特征的修正,可以提高最终的分割效果。

1.3 改进的 U - Net 神经网络

U - Net 架构对于复杂度并不高的基础模型能

够达到很高的精度,且能在精度与复杂度之间达到平衡,但是遥感影像背景复杂,所包含的地物种类丰富,光谱范围广泛,直接采用 U-Net 网络构架并不能有效地提取复杂的遥感影像的像元特征。鉴于 U-Net 网络训练过程中收敛速度快,为满足包含多语义信息的高分遥感影像信息提取,本文使用上述 U-Net 网络作为基本构架对其进行改进,具体过程如下:

- 1)该网络结构与主流的图像分割网络一样采用对称的 encoder to decoder 结构,使用随机位移矢量算法使训练集产生弹性形变,增加训练集特征。
- 2)U-Net 网络进行部分卷积优化,从训练图像和标签中提取成对的补丁,以增加训练数据集的大小,并使训练更加稳健。
- 3)在训练过程中每个小批包含 16 个补丁,在每一个时代的每一次迭代中,都会从图像中的随机位置提取出上千个小批量补丁。
- 4)采用的 U-Net 网络的 up-sampling 依然有大量的通道,这使得网络将上下文信息向恢复 feature map 的更高分辨率传播。
- 5)设置初始学习率为 0.005,并使用 L2 范数控制学习率的渐变曲线。

2 实验及分析

DTRF-U-Net 网络模型整体方案和训练流程如图 1 所示,该方案的实施主要分为数据预处理、模型训练和结果分析 3 个模块。为验证本文方法的提取效果,分别采用 U-Net 和 DTRF-U-Net 网络模型分别对研究区遥感影像中的建筑物进行提取,并做对比分析。

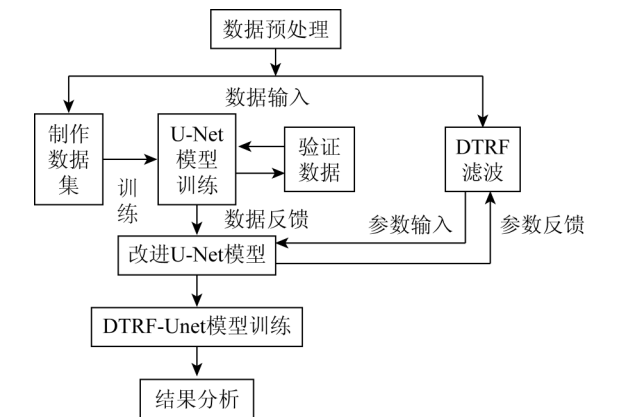


图 1 实验流程图
Fig. 1 Flow chart of the experiment

2.1 实验数据及数据集制作

实验区定址于辽宁省盘锦市紧接渤海湾地区,平均海拔约 30 m,中心经纬度分别为 E121. 16°, N41. 12°。实验区的地貌主要类型是大量冲海积

平原和潮滩,地貌类型相对单一,地势平坦,土地类型主要为建筑用地和耕地。选用 2019 年 9 月 29 日高分二号影像(图 2)。高分二号卫星搭载 2 台高分分辨率 1 m 全色、4 m 多光谱相机,具有米级空间分辨率。首先对高分二号影像的全色波段和多光谱数据进行预处理,包括正射校正、数据融合和裁剪等操作^[26]。因 U-Net 模型采用数据增强得方式可有效利用训练数据,加之考虑到 RTX2060 显卡计算能力和影像质量因素,故在较多建筑物的影像中心区域,选取尺寸为 1 829 像素×1 410 像素作为验证数据集,选取尺寸为 954 像素×936 像素作为训练集,选取尺寸为 1 718 像素×1 303 像素作为测试集。最后对数据集进行人工标注,使之成为二通道数据,即建筑物标注为 1,其他标注为 0,与数据集共同输入 DTRF-U-Net 模型计算。



图 2 实验样本
Fig. 2 Experimental samples

为充分发挥深度学习的优势,所以对训练集采用面向对象分类的方法,充分利用遥感影像的空间信息、光谱、纹理、相邻关系、形状等特征进行分割分类,采用此方法标注影像得到地表真实地物(ground truth);然后对数据集进行数据增强,并根据数据集图像的均值和方差对所有图像进行归一化,将数据集输入集成了多尺度交叉训练、多重损失计算、引入一阶动量的随机梯度下降(stochastic gradient descent + momentum, SGDM)优化的神经网络模型 U-Net 中进行训练。在训练过程中,经数据预处理后每次选取的训练样本大小为 256×256,训练和验证的数据集均包含与相应真实地物对应的标签数据集;然后将训练集图像输入到 U-Net 网络中,进行充分训练后得到优化的网络模型,最后将测试图像输入训练好的模型中,得到预测结果,根据滤波前后的训练结果计算分析。

2.2 实验结果与分析

经改进后的影像预处理的过程中需要首先使用

DTRF 滤波器进行处理,为此,需要确定滤波器的信号空域标准差 δ_s 和值域标准差 δ_r 两个平滑参数,不同的平滑参数会导致不同的结果。经过如下过程确定合适的取值区间,上行 δ_s 设置为 30, δ_r 分别设置为 0.2,0.4,0.6; 下行设置 δ_s 为 60, δ_r 分别设置为 0.2,0.4,0.6。图 3 以图 2 中红框区域为例展示不同平滑参数的滤波效果,其中图 3(a)为原始数据。可以看出,在 δ_s 分别为 30 和 60 的情况下,当 δ_r 为

0.6 时图像发生明显的涂抹效应,当 δ_s 为 60 时会比 δ_s 为 30 时涂抹效应更严重。但是当滤波窗口非极大值时,滤波后图像与输入图像鲜有差异,同时当超过某一范围后,滤波结果将不会发生太大的变化。也就是说明当参数设置恰当时,DTRF 滤波不仅可以有效去除图像中的噪声、细节纹理等信息,同时可以增强图像中空间结构信息。

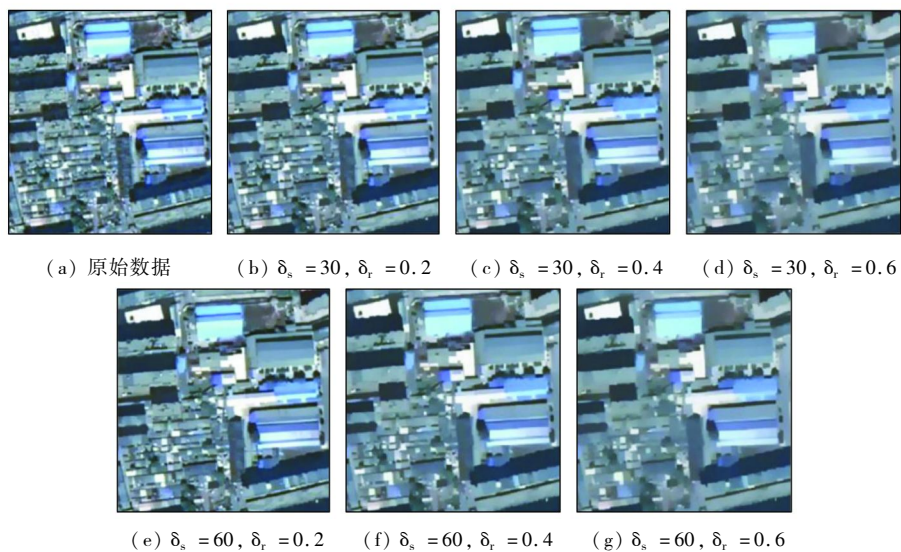


图 3 不同参数滤波效果

Fig. 3 Filtering effect with different parameters

图 3 已表明不同的 δ_s 和 δ_r 会带来不同的数据处理精度,为此,进一步取在 30 ~ 60 区间按步长 15 遍历 δ_s 值, δ_r 按步长 0.1 遍历区间 0.2 ~ 0.8,分别

计算精度 (precision)、标准差 (standard deviation, SD)、结构相似性 (structural similarity, SSIM) 和平均交并比 (MIoU) 4 个指标来衡量提取质量 (图 4)。

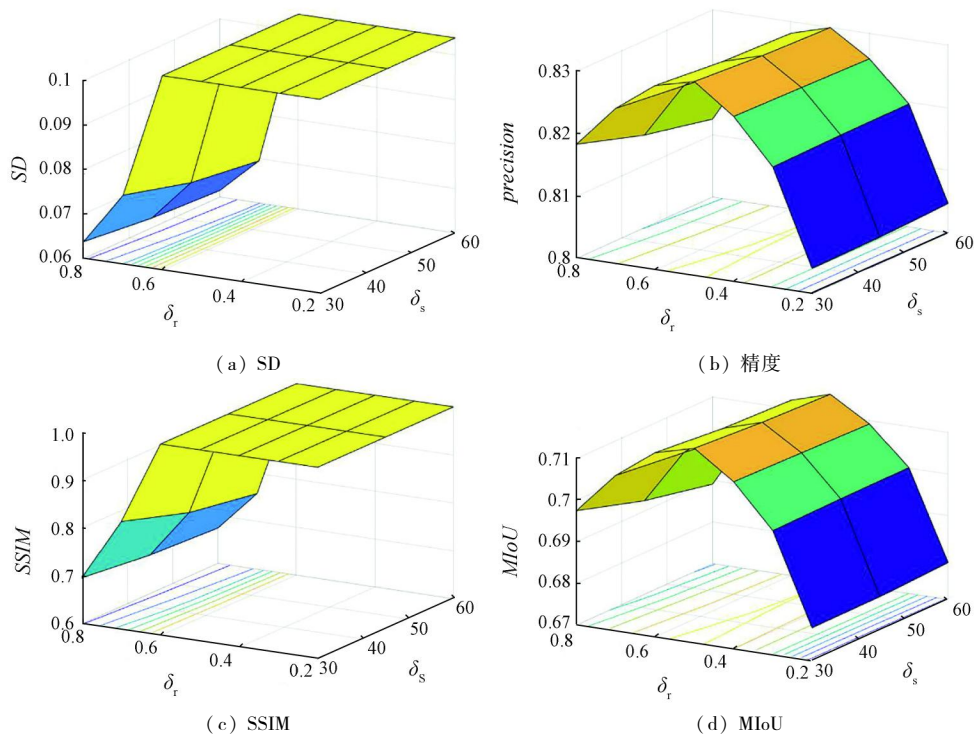


图 4 不同参数对 DTRF - UNET 输出数据影响

Fig. 4 Influence of different parameters on DTRF - Unet output data

SD 可以反映出图像像素值与均值的离散程度,SD 值大小与图像质量呈正相关;SSIM 表现出处理前后数据的结构相似性,是基于图像的亮度、对比度和结构 3 个要素来评定,值越大代表与原始数据相比保留细节越多,信息保留程度越高;MIoU 也是在深度学习中衡量分割精度的重要指标,即在每个类别上计算交并比取均值,与分割精度同样保持正相关性。整体分类精度最终结果列于表 1。

表 1 不同参数下 DTRF - Unet 整体分类精度
Tab.1 Overall classification accuracy of DTRF -

Unet under different parameters		(%)						
δ_s	δ_r							
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	
30	80.41	81.94	82.65	83.12	82.96	82.50	81.83	
45	80.42	81.96	82.63	83.04	82.86	82.37	81.49	
60	80.46	81.97	82.60	83.00	82.82	82.25	81.25	

可以看出,当 δ_r 在 0.6 ~ 0.8 之间 SD 最小(图 4(a));当 δ_r 在 0.5 ~ 0.6 之间精度较高,最低处在 δ_r 为 0.2 处为 80.41%(图 4(b)); δ_r 在 0.2 ~ 0.6 之间 SSIM 没有发生明显变化, δ_r 为 0.7 ~ 0.8 时 SSIM 开始下滑,意味着逐渐与原图像异质化(图 4(c));MIoU 用于测量真实和预测之间的相关度,相关度越高,该值越高,一般来说大于 0.5 可以认为结果较好,图 4(d)中最低值在 δ_r 处于 0.2 时取到,但结果均大于 0.67,在 δ_r 为 0.4 时取到极大值。综合图 4 可以看出:整体质量保持在 δ_r 为 0.5 时精度达到最高, δ_s 对分类质量没有明显影响,基本处于同一等

高线,经滤波后的整体分类精度近似保留两位小数后均保持在 80% 以上,识别效果较好。根据表 1 可知在 δ_s 为 30, δ_r 为 0.5 时取得最佳精度 83.12%,较滤波前 U - Net 分类精度的 75.99% 提高了 7.13 百分点,验证了 DTRF - Unet 方法作为改进的 U - Net 深度学习网络处理图像的有效性。

图 5 以图 2 中黄色框区域为例对比改进前后的提取效果。样本区选择在城市群中,分为建筑物和其他类别两类,深蓝色为建筑物类,天蓝色为其他类。利用原始 U - Net 深度学习方法对样本数据进行提取,从图 5(a) 提取建筑物的结果可以看出 U - net 在城市群处理过程中提取精度较高,建筑物轮廓较明显,部分道路线也对城市集群有明显的分割效果,且基本没有细碎的噪声斑块,东北方向的林地也被正确地划分到了其他类中,但是整体精度只有 75.99%。最后对 DTRF - Unet 的 21 个结果汇总分析极化结果之间的差异如图 5 所示,发现在样本区的东北方向和中部偏南方向箭头标注方向有明显差异,图 5(b) 较图 5(a) 和图 5(c) 更接近地面真实情况,图 5(b) 将部分建筑用地误判为其他类,与上文验证精度相一致,但是在部分地区还是存在明显的漏分错分误差,可能受到训练样本数据大小和影像阴影区域较大影响。图 5(b) 通过选取适当的参数有效剔除了噪声,使建筑物边界更加平滑,预测结果得到进一步优化。

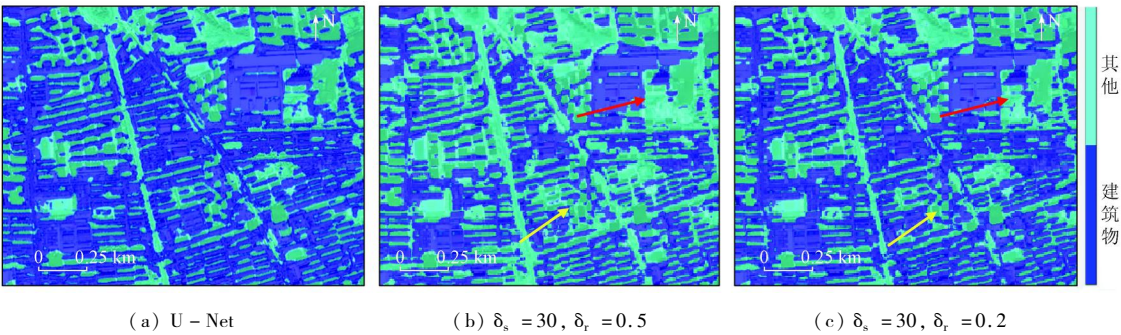


图 5 改进前后结果对比
Fig.5 Comparison of results before and after improvement

3 结论

1) 针对常规 U - net 神经网络模型在高分影像最大池化采样中,可能丢失一些细节信息,尤其是图像边缘特征的损失,从而导致建筑物提取不全或边缘模糊问题,提出了一种基于常规 U - Net 的改进的智能建筑物识别提取方法,即 DTRF - Unet 模型。

2) 该模型能很好地将低层高分辨率信息与高层抽象的语义信息融合减小建筑物细节特征损失,既增强了网络的泛化表达能力,又提高了建筑物提取精度。通过与原始结果对比发现,该方法可以近似完整地将建筑物自动分割出来,具有高效性和可实施性。

3) 在滤波预处理时,采用以一定步长遍历某段区间,观察滤波效果进而确定 DTRF 两个平滑参数的方法虽然效率较低,但是在目前尚缺乏实质高效

的自适应方法条件下,不失为一种有益尝试,为此类研究提供一定参考。但是受计算能力限制,模型分割不够细致,时间性能还有待提升,同时该实验对于部分阴影区域的识别精度有待商榷,这将是下一步研究的重点内容。

参考文献(References):

- [1] 王伟鹏,戴声奎. 基于域变换递归滤波的雾天图像复原[J]. 计算机应用研究,2015,32(2):638-640.
Wang W P, Dai S K. Haze image restoration using domain transform recursive filter[J]. Computer Application Research, 2015, 32(2):638-640.
- [2] 彭海涛,柯长青. 基于多层分割的面向对象遥感影像分类方法研究[J]. 遥感技术与应用,2010,25(1):149-154.
Peng H T, Ke C Q. Study on object-oriented remote sensing image classification based on multi levels segmentation[J]. Remote Sensing Technology and Application, 2010, 25(1):149-154.
- [3] Camps-Valls G, Gomez-Chova L, Munoz-Mari J, et al. Composite kernels for hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2006, 3(1):93-97.
- [4] 沈毅. 高分辨率遥感图像处理若干关键技术研究[D]. 南京:南京航空航天大学,2012.
Shen Y. Research on some key technologies of high-resolution remote sensing image processing[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2012.
- [5] 桂海刚,冯文卿,李文卓,等. 多时相遥感影像变化检测方法综述[J]. 武汉大学学报(信息科学版),2018,43(12):1885-1898.
Sui H G, Feng W Q, Li W Z, et al. Review of change detection methods for multi-temporal remote sensing imagery[J]. Geomatics and Information Science of Wuhan University, 2018, 43(12):1885-1898.
- [6] Baltsavias E P. Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2004, 58(3/4):129-151.
- [7] Alarcon-Ramirez A, Rwehangira M R, Chouikha M F, et al. A new methodology based on level sets for target detection in hyperspectral images[J]. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(9):5385-5396.
- [8] Wei G, Wen Y, Haijian Z, et al. Geospatial object detection in high resolution satellite images based on multi-scale convolutional neural network[J]. Remote Sensing, 2018, 10(1):131.
- [9] 方旭,王光辉,杨化超,等. 结合均值漂移分割与全卷积神经网络的高分辨遥感影像分类[J]. 激光与光电子学进展,2018,55(2):022802.
Fang X, Wang G H, Yang H C, et al. High resolution remote sensing image classification combined with mean shift segmentation and full convolutional neural network[J]. Advances in Laser and Optoelectronics, 2018, 55(2):022802.
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems, 2012, 60(2):1097-1105.
- [11] Mou L, Ghamisi P, Zhu X X. Deep recurrent neural networks for hyperspectral image classification[J]. IEEE Transaction Geoscience and Remote Sensing, 2017, 55(7):3639-3655.
- [12] 陈睿敏,孙胜利. 基于深度学习的红外遥感信息自动提取[J]. 红外,2017,38(8):37-43.
Chen R M, Sun S L. Automatic extraction of infrared remote sensing information based on deep learning[J]. Infrared, 2017, 38(8):37-43.
- [13] 王宁,程家骅,张寒野,等. U-net模型在高分辨率遥感影像水体提取中的应用[J]. 国土资源遥感,2020,32(1):35-42. doi:10.6046/gtzyyg.2020.01.06.
Wang N, Cheng J H, Zhang H Y, et al. Application of U-net model to water extraction with high resolution remote sensing data[J]. Remote Sensing for Land and Resources, 2020, 32(1):35-42. doi:10.6046/gtzyyg.2020.01.06.
- [14] 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 Y S, Zhen H W, Luo G P, et al. Extraction and counting of Populus Euphratica crown using UAV images integrated with U-Net method[J]. Remote Sensing Technology and Application, 2019, 34(5):939-949.
- [15] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015, 79(11):3431-3440.
- [16] 苏健民,杨岚心,景维鹏. 基于U-Net的高分辨率遥感图像语义分割方法[J]. 计算机工程与应用,2019,55(7):207-213.
Su J M, Yang L X, Jing W P. U-Net based semantic segmentation method for high resolution remote sensing image[J]. Computer Engineering and Applications, 2019, 55(7):207-213.
- [17] 张慧,王坤峰,王飞跃. 深度学习在目标视觉检测中的应用进展与展望[J]. 自动化学报,2017,43(8):1289-1305.
Zhang H, Wang K F, Wang F Y. Advances and perspectives on applications of deep learning in visual object detection[J]. Journal of Automation, 2017, 43(8):1289-1305.
- [18] 伍广明,陈奇,Shibasaki R,等. 基于U型卷积神经网络的航空影像建筑物检测[J]. 测绘学报,2018,47(6):864-872.
Wu G M, Chen Q, Shibasaki R, et al. High precision building detection from aerial imagery using a U-Net like convolutional architecture[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(6):864-872.
- [19] 向轩琳. 基于域变换递归滤波的高光谱图像分类方法研究[D]. 长沙:湖南大学,2018.
Xiang X L. Research on hyperspectral image classification method based on domain transform recursive filter[D]. Changsha: Hunan University, 2018.
- [20] Gastal E S L, Oliveira M M. Domain transform for edge-aware image and video processing[C]//ACM Transactions on Graphics. 2011, 30(4):1-12.
- [21] Kang X D, Duan P H, Li S T. Hyperspectral image visualization with edge-preserving filtering and principal component analysis[J]. Information Fusion, 2020, 57:130-143.
- [22] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing

and Computer – Assisted Intervention. Springer,2015 ;234 – 241.

[23] Badrinarayanan V ,Kendall A ,Cipolla R. SegNet; A deep convolutional encoder – decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence , 2017 ,39(12) ;2481 – 2495.

[24] 张浩然 ,赵江洪 ,张晓光. 利用 U – net 网络的高分遥感影像建筑提取方法[J]. 遥感信息 ,2020 ,35(3) ;143 – 150.

Zhang H R ,Zhao J H ,Zhang X G. High resolution remote sensing image building extraction method using U – net network[J]. Remote Sensing Information ,2020 ,35(3) ;143 – 150.

[25] Li R ,Liu W ,Yang L ,et al. DeepUNet; A deep fully convolutional network for pixel – level sea – land segmentation[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing ,2018 ,11(11) ;3954 – 3962.

[26] 郝建亭 ,杨武年 ,李玉霞 ,等. 基于 FLAASH 的多光谱影像大气校正应用研究[J]. 遥感信息 ,2008(1) ;78 – 81.

Hao J T ,Yang W N ,Li Y X ,et al. Atmospheric correction of multi – spectral imagery ASTER [J]. Remote Sensing Information ,2008(1) ;78 – 81.

Research on building cluster identification based on improved U – Net

WU Yu¹ , ZHANG Jun¹ , LI Yixu² , HUANG Kangyu¹

(1. School of Mining , Guizhou University , Guiyang 550025 , China ; 2. College of Agriculture , Guizhou University , Guiyang 550025 , China)

Abstract: Aiming at tackling the problem that some edge features of buildings are easily blurred or lost in the extraction of buildings with high resolution image by U – Net , this paper proposes an optimized building extraction method , which firstly enhances the edge of buildings with high resolution image and simultaneously improves the partial convolution process of U – Net. Specific process is as follows : Firstly , the domain change recursive filtering method is used to enhance the edge of the building , and the enhanced image is input into U – Net neural network results for training. To make full use of the rich details characteristics of the buildings on the GF – 2 images , the authors tried to extract pairs from training images and label patch on the basis of the original U – Net structure and in the process of coding decoding , so as to increase the training data. These patches further strengthened the positive and negative deep learning of high – dimensional feature for buildings , thus successfully realizing building image segmentation. In this paper , the experimental results of the extraction of GF – 2 image buildings in Panjin City of Liaoning Province adjacent to Bohai Bay on September 29 , 2017 show that the overall classification accuracy of the buildings detected by U – Net is 75.99% for the shaded and unsatisfied area sample data , and the maximum overall classification accuracy of this method can reach 83.12% , which is 7.13 percentage higher than that of the original U – Net network. It is proved that the U – NET model combined with domain change recursive filtering is effective.

Keywords: deep learning ; domain change recursive filtering ; U – Net ; edge enhancement ; building extraction
(责任编辑: 张 仙)