

文章编号: 2095-2163(2023)09-0051-08

中图分类号: TP751;TP389.1

文献标志码: A

基于U型残差网络的遥感图像道路提取方法研究

李梓瑜, 王大东, 于晓鹏

(吉林师范大学 数学与计算机学院, 吉林 四平 136000)

摘要: 针对遥感图像道路提取时大量建筑物、植被等干扰信息遮挡, 导致出现错分、漏分、边缘信息恢复不完整等问题, 提出一种基于 Adaptive Mixup 操作的深度残差分割网络。首先, 使用预训练的 ResNet101 作为网络的编码器, 有效地保持其特征表达能力; 其次, 引入 ASPP 模块提取图像的多尺度特征信息, 在特征融合过程中加入 Adaptive Mixup 操作, 对特征信息进行自适应地动态融合; 最后, 使用转置卷积与 Pixelshuffle 结合的上采样方法, 对缩小后的特征图进行有效放大。通过使用两个高分辨率遥感数据集 Massachusetts Roads Dataset 和 DeepGlobe Roads Dataset 的实验结果表明, 所提方法最高可达 92.28% 的精确率值和 86.97% 的 F1 值, 改进后的网络在各项评价指标上均优于 FCN、SegNet 等经典语义分割网络, 能够有效对道路信息进行提取, 具备一定的实用价值。

关键词: 遥感图像; 道路提取; Adaptive Mixup 操作; 残差分割网络; Pixelshuffle 操作

Research on road extraction method of remote sensing image based on U-shaped residual network

LI Ziyu, WANG Dadong, YU Xiaopeng

(College of Mathematics and Computer Science, Jilin Normal University, Siping Jilin 136000, China)

[Abstract] Aiming at the problem of too much feature information in the process, and low segmentation accuracy caused by a large number of buildings, vegetation and other interference information, a deep residual segmentation network based on adaptive mixup operation is proposed. Firstly, the pre-trained Resnet 101 is used as the encoder of the network to effectively maintain its feature expression ability and solve the problem of gradient vanishing or gradient explosion to a certain extent; Secondly, the ASPP module is introduced to extract the multi-scale feature information of the image; Then, adaptive mixup operation is added in the process of feature fusion to both adaptively and dynamically fuse the feature information; Finally, in the information fusion of upsampling and shallow features, the pixelshuffle operation is used to replace the original upsampling method to effectively enlarge the reduced feature map. Experiments using two high-resolution remote sensing data sets Massachusetts Roads Dataset and DeepGlobe Roads Dataset show that the proposed method can reach up to 92.28% precision rate value and 86.97% F1 value. The improved network outperforms the classical semantic segmentation networks such as FCN and SegNet in various evaluation indexes, which can effectively extract road information, thus, it has certain practical application value.

[Key words] key remote sensing image; road extraction; adaptive mixup operation; residual segmentation network; pixel shuffle operation

0 引言

道路是城市规划、地理信息系统更新和交通导航等许多重要领域的主干和基础设施。近年来,随着卫星观测技术的飞速发展,使得高分辨率遥感图像的道路提取问题成为人们关注的焦点。然而,通过传统人工标注提取信息的方式来分割道路,不仅费

时费力,且只能提取到图像表层信息,所得的分割结果也存在较大误差^[1]。因此,针对遥感道路提取的自动化处理方法就显得尤为重要。

利用高分辨率遥感图像进行道路分割一直是遥感领域研究的重难点,遥感图像从空中俯拍包含的物体繁多,同类物体的颜色、纹理、大小极其相似,且植被、建筑等影响因素的遮挡给分割任务带来巨大

基金项目: 国家自然科学基金(61773009); 教育部科技发展中心产学研创新“智融兴教”基金(2018A02002)。

作者简介: 李梓瑜(1998-),女,硕士研究生,主要研究方向:深度学习、遥感图像分割方法;王大东(1971-),男,博士,副教授,主要研究方向:数字图像处理与深度学习;于晓鹏(1964-),男,学士,教授,主要研究方向:深度学习与数值模拟研究。

通讯作者: 王大东 Email:jlmspydr@163.com; 于晓鹏 Email:yyxpp@jlnu.edu.cn

收稿日期: 2022-10-05

的难度^[2]。2015年,Long等人^[3]提出了图像分割领域具有开创性意义的全卷积网络(Fully Convolutional Network, FCN),该方法利用反卷积替换了卷积神经网络(Convolutional Neural Networks, CNN)中的全连接,实现了端到端的网络训练。这种基于像素的方法相较于CNN和传统人工提取虽更为高效,但对图像信息保留不够完整。同年,Ronneberger等人^[4]提出的U-Net网络实现了多尺度信息的融合,因其性能优越且训练速度较快,目前广泛应用于图像分割领域。随着深度学习技术在计算机视觉领域的发展进步,众多学者经多年研究对深度学习方法做出改进,旨在提高遥感图像道路提取的精确度。2016年,He等人^[5]提出残差网络(ResNet)对更深层次信息进行提取,在增加网络深度的同时,提高网络训练结果的精确度。2018年,Zhang等人^[6]受残差网络启发,将U-Net与ResNet进行结合,提出ResUNet网络用于道路特征提取,简化了深层网络的训练,并充分利用跳跃连接实现模型内部的信息传递,在图像分割领域取得良好的效果。He等人^[7]将空间金字塔池化模块(Atrous Spatial Pyramid Pooling, ASPP)^[8]与编码-解码网络结构相结合,实现了对道路特征更加精细的提取。Zhou等人^[9]基于LinkNet^[10]和空洞卷积(Dilated convolution)开发了一个名为D-LinkNet的Encoder-Decoder网络,借助更大的感受野,融合提取到的低级、高级语义特征,最终赢得了2018年国际计算机视觉与模式识别会议中DeepGlobe^[11]道路提取挑战赛(CVPR DeepGlobe Road Extraction Challenge)的第一名。Yang等人^[12]在U-Net网络的基础上设计了一个循环卷积神经网络模块,能够更好的提取空间上下文信息,实现道路提取。Han等人^[13]提出基于带孔卷积改进的残差网络和基于密集连接改进的空洞空间卷积池化金字塔模块的图像分割网络,利用遥感图像不同尺度特征信息,有效提高城市地区的图像分割效果。Xiao等人^[14]基于残差网络、ASPP和门控卷积开发了Gated-ResNet网络,使得提取的道路信息更加完整,在图像分割方面取得不错的效果。Chen等人^[15]将残差网络与非对称卷积块进行结合,提出一种编码-解码器结构的AFU-Net网络,对不同层次信息进行多尺度融合,使得图像边缘信息的提取更加清晰明了。

在语义分割任务中,低级特征(如:边缘和轮廓)可以在卷积神经网络的浅层中捕获,而随着网络深度的增加,浅层特征逐渐退化,且传统的低级特

征与高级特征往往采用固定比例的方式进行融合,但通常会出现细节特征与语义特征丢失的现象。

综上所述,虽然在遥感图像道路提取任务中取得良好的分割效果,但是在深层次语义特征和浅层纹理特征的提取融合方面表现一般,导致地物信息复杂且道路遮挡严重的遥感图像提取效果不佳。为了解决上述问题,本文受U形网络结构、残差连接、像素重组(Pixelshuffle)、多尺度特征融合和自适应混合(Adaptive Mixup)等操作的启发,提出一种使用超参数自适应操作,能够调节高级特征与低级特征混合比例的残差分割网络AMP-ResUNet(ASPP+Mixup+Pixelshuffle-ResUNet),以提升遥感图像道路分割的精度及完整度。

1 AMP-ResUNet网络工作原理

如图1所示,AMP-ResUNet是一种端到端的网络结构模型。首先,在编码器中使用预训练的ResNet101网络替换掉原始U-Net网络中的下采样部分,在有效保持其特征表达能力的同时,在一定程度上解决因网络层数加深而导致的梯度消失或梯度爆炸问题。其次,在编码器与解码器的衔接部分引入空洞空间金字塔池化模块,并对模块中扩张率进行改进,避免由于扩张率过大带来的模型退化问题,实现对图像的多尺度特征提取。在特征融合过程中加入Adaptive Mixup操作,使得来自下采样部分的浅层特征信息自适应地从上采样部分流向高级特征,对特征信息进行动态融合。最后,使用转置卷积与Pixelshuffle操作结合的方式作为网络的解码部分,在上采样的最后一层使用Pixelshuffle操作,替换传统基于数学的双线性插值和填充零操作,对缩小后的特征图进行有效放大,提升网络的运行效率和整体性能。

1.1 AMP-ResUNet网络编码结构

众所周知,在训练神经网络模型的过程中,随着网络层数不断加深,会出现网络“退化”现象,ResNet网络中残差模块的提出,能够使深层网络训练出的模型效果优于浅层网络,有效缓解了这一现象。该模块的引入不仅增加了神经网络的深度,还能有效保持其特征表达能力,在一定程度上解决了因网络层数加深而引发的梯度消失或梯度爆炸问题。模块中每个残差单元可表示为

$$x_{j+1} = x_j + F(x_j, w_j) \quad (1)$$

式中: x_j 代表该层网络的输入信息, x_{j+1} 则代表输出信息, w_j 表示该层待学习的参数。

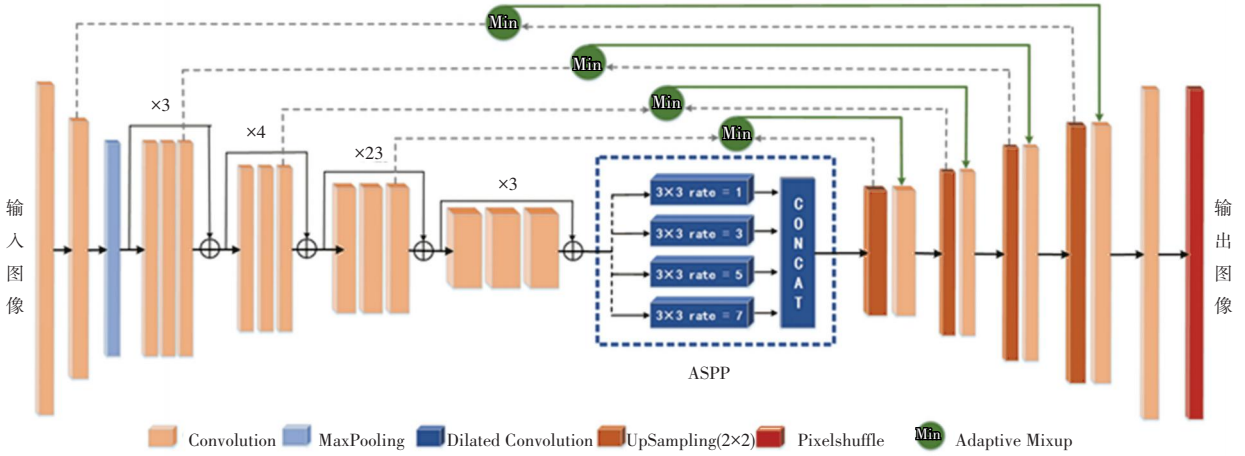


图 1 AMP-ResUNet 网络结构

Fig. 1 Network structure of the AMP-ResUNet

将式(1)进行递归运算,得到任意深层单元特征表示如式(2):

$$x_j = x_j + \sum_{i=j}^{J-1} F(x_i, w_i) \quad (2)$$

本文使用 Pytorch 官方提供的预训练 ResNet101 网络作为编码器来提升特征提取的效果,该网络以 VGG 网络为基础,基于短路机制添加残差学习模块搭建。为适应本文模型结构,去掉原始 ResNet101 网络中的平均池化层和分类层,其具体结构如图 2 所示。

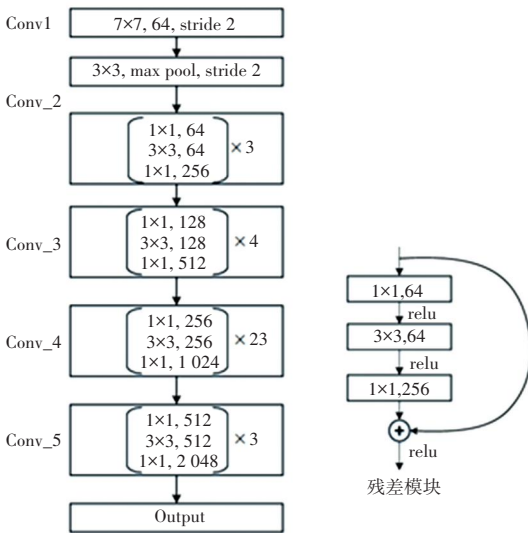


图 2 ResNet101 网络结构图

Fig. 2 Network structure of the ResNet101

1.2 改进的 ASPP 模块

空洞空间金字塔池化 (Atrous Spatial Pyramid Pooling, ASPP) 在 DeepLab 等网络中被广泛应用。该模块中的空洞卷积,可以有效地增加神经元的感受野,对所给定的输入以不同扩张率

行并行采样,从多尺度捕捉图像的特征信息,模块中每个空洞卷积可表示为式(3):

$$y[i] = \sum_{k=1}^K x[i + r \cdot k] \omega[k] \quad (3)$$

式中: i 表示每个遥感图像的像素点, x 为空洞卷积的输入, y 为网络输出, $\omega[k]$ 代表过滤器的长度为 k , r 是步长的扩张率,通过给定 r 的不同数值来改变过滤器的感受野。

原始的 ASPP 是由扩张率为 1、6、12、18 的 4 个卷积核组成的,但是当扩张率过大时,会产生无意义的权重,导致有效权重的卷积核数量减少。因此,本文对 ASPP 模块的扩张率进行改进。具体结构如图 3 所示。

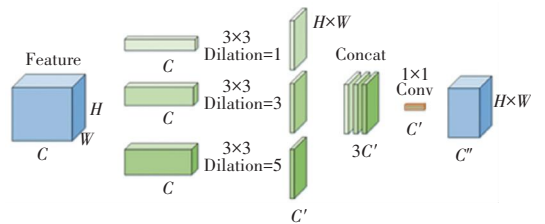


图 3 ASPP 模块

Fig. 3 The module of ASPP

本文使用空洞率为 1、3、5 的扩张卷积核,其对应的感受野大小分别为 3×3 、 7×7 、 11×11 ,之后将这 3 张特征图进行通道维度的拼接,最后进行一次卷积核大小为 1×1 的卷积操作,对通道数进行压缩。由于本文对扩张率大小进行改进并选择了合适的采样率,不存在因扩张率过大导致的模型退化问题,因此取消原始 ASPP 模块中的池化层。

1.3 Adaptive Mixup 操作

由于低级特征通常可以在卷积神经网络的浅层

中被捕获,但随着网络深度的增加,浅层特征逐渐退化。为解决这个问题,已有很多研究通过添加或串联跳跃连接,将浅层特征与深层特征进行结合,辅助图像重构。虽然跳跃连接在一定程度上缓解了细节丢失问题,但仍然存在一定的改进空间。2021年, Wu 等人^[16]提出自适应混合操作(Adaptive Mixup Operation),对上采样层和下采样层之间的特征进行动态融合,通过改进层间的信息流动方式,提升特征融合的效果。该操作结构如图4所示。

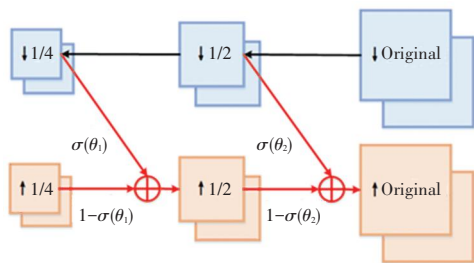


图4 Adaptive Mixup 操作

Fig. 4 Adaptive Mixup operation

由图中可见,第一行和第二行分别进行上采样和下采样操作。传统的特征融合是高级特征与低级特征一比一进行融合,而 Adaptive Mixup 使得来自下采样部分的浅层特征信息,自适应地从上采样部分流向高级特征,混合运算的最终输出可以表示为:

$$f_{\downarrow 2} = \text{Mix}(f_{\downarrow 1}, f_{\uparrow 1}) = \sigma(\theta_1) * f_{\downarrow 1} + (1 - \sigma(\theta_1)) * f_{\uparrow 1}$$

$$f_{\uparrow 1} = \text{Mix}(f_{\downarrow 2}, f_{\downarrow 2}) = \sigma(\theta_2) * f_{\downarrow 2} + (1 - \sigma(\theta_2)) * f_{\downarrow 2}$$

$$(4)$$

式中: $f_{\downarrow i}$ 和 $f_{\uparrow i}$ 分别是来自第*i*个下采样层和上采样层的特征映射, f_{\uparrow} 是最终输出。 $\sigma(\theta)_i$, $i=1,2$ 是融合来自第*i*下采样层和第*i*上采样层输入的第一个可学习因素,其值由参数 θ_i 上的符号算子 σ 决定。在训练过程中,可以对这两个因素进行有效率的学习,使其性能和效果远好于常量因素。

1.4 Pixelshuffle 操作

AMP-ResUNet 使用转置卷积与 PixelShuffle 操作相结合的方式上进行上采样,但由于转置卷积生成的图像会出现边缘生硬和不真实的情况,因此在解码器部分的最后一层将转置卷积替换为 PixelShuffle 上采样方法,对缩小后的特征图进行有效的放大,使其边缘恢复更加完整,提升网络的鲁棒性和准确性。PixelShuffle 的主要功能是将低分辨率的特征图,通过卷积和多通道间的重组得到高分辨率的特征图。文献[17]中提出亚像素卷积层(sub-pixel convolutional layer)的方法来扩大特征图,具体结构如图5所示。

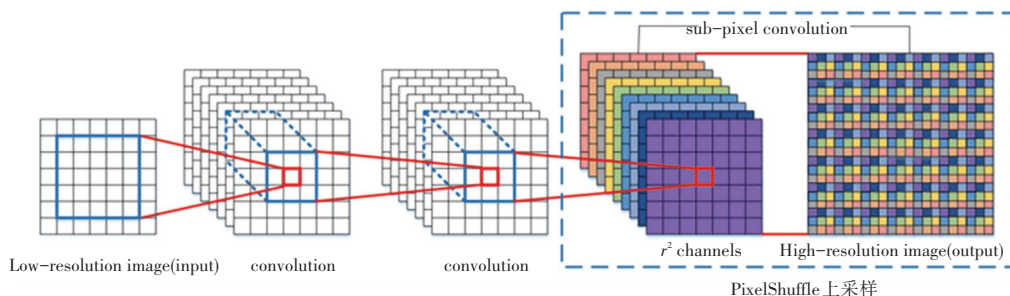


图5 亚像素卷积神经网络(ESPCN)

Fig. 5 Subpixel convolutional neural network (ESPCN)

如图5所示,网络的输入是原始低分辨率图像,通过两个卷积层以后,特征图像与输入图像大小一样,特征通道变为 r^2 。再将每个像素的 r^2 个通道重新排列成一个 $r \times r$ 的区域,对应于高分辨率图像中的一个 $r \times r$ 大小的子块,从而大小为 $r^2 \times H \times W$ 的特征图像被重新排列成 $1 \times rH \times rW$ 大小的高分辨率图像。通过使用 sub-pixel convolution 方法,图像从低分辨率到高分辨率放大的过程中,可以被自动学习到的插值函数被隐含地包含在卷积层中。由于在低分辨率图像上进行卷积运算,只在最后一层对图像大小进行变换,因此所需处理时间较短,提高了

模型的运算效率。图中彩色部分从 r^2 channels \rightarrow High-resolution image 的示意过程即为 PixelShuffle。因此,PixelShuffle 可以看成是一个特殊的重组操作,通过将通道维度的像素向长宽维度搬移来实现上采样。

1.5 损失函数

损失函数是用来评估模型训练效果的一个标准,简单来说就是用来表现预测值与实际数据的差距程度,损失函数值越小,代表其模型的鲁棒性就越好。本文使用的是 MS-SSIM(多尺度结构相似)损失函数和 Dice 损失函数。MS-SSIM 损失函数的公

式如式(5):

$$MS - SSIMLoss = 1 - \prod_{m=1}^M \left(\frac{2\mu_p\mu_g + C_1}{\mu_p^2 + \mu_g^2 + C_1} \right)^{\beta_m} \left(\frac{2\sigma_{pg} + C_2}{\sigma_p^2 + \sigma_g^2 + C_2} \right)^{\gamma_m} \quad (5)$$

式中: M 表示尺度的总数量, $\mu_p, \mu_g, \sigma_p, \sigma_g$ 和 σ_{pg} 分别表示预测图片与地面真实值的均值、标准差和协方差。 β_m, γ_m 为两者间的相对重要性。 C_1, C_2 防止除数为0。 $MS - SSIM$ 损失函数赋予了模糊边缘更高的权重, 区域分布差异越大, $MS - SSIM$ 值越高。

$Dice$ 损失函数是一种计算样本之间相似度的度量函数, 是把一个类别中的所有像素看做一个整体进行计算, 在一定程度上解决了正负样本不均衡的问题, 且收敛速度很快。 $Dice$ 损失函数的公式如式(6):

$$DiceLoss = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (6)$$

式中: X 代表地面真实道路面积的区域, Y 代表预测道路面积的区域, 取值范围在 0~1 之间。

本文使用的总损失函数是将 $MS - SSIM$ 损失函数和 $Dice$ 损失函数 1:1 进行相加, 具体公式如式(7):

$$Loss = MS - SSIMLoss + DiceLoss \quad (7)$$

2 实验

2.1 实验环境

本文实验代码基于 Pytorch 框架构建, 编译环境为 Python3.8, 操作系统为 64 GB 内存的 Ubuntu 20.04 LTS。硬件配置 GPU 型号为 Intel Xeon Gold 5215@ 2.50 GHz, 显卡为 NVIDIA GeForce RTX 2080Ti。

2.2 数据集

本文选取美国马萨诸塞州道路数据集 (Massachusetts Roads Dataset) 和 DeepGlobe 遥感图像道路提取数据集作为实验数据。Massachusetts 道路数据集是目前最大的遥感影像道路数据集, 覆盖面积超过 2 600 km^2 , 覆盖地物信息主要包含城市、农村、郊区的道路分布。数据集中共有 1 171 张大小为 1 500×1 500 像素的遥感图像, 地面分辨率约为 1m/像素, 包含 1 108 张训练集、49 张测试集和 14 张验证集, 每组数据集图像如图 6 所示。

由于东南亚地区与美国马萨诸塞州道路场景存在较大差距, 则选取 DeepGlobe 道路数据集, 来验证算法的可行性与泛化性。该数据集中包含 6 226 张

1 024×1 024 像素大小的训练图像及其对应的标签, 每张图像都是由 DigitalGlobe 卫星采集的地面分辨率为 0.5 m/pixel 的 RGB 图像, 覆盖范围包括东南亚多个国家的郊区、雨林等不同场景的道路分布。将数据集随机分为 5 800 张训练集、178 张测试集和 248 张验证集。其中, 便签中道路信息与非道路信息像素分别为 255 和 0, 是与输入图像有着相同尺寸的灰度二值图像。每组数据集图像如图 7 所示。

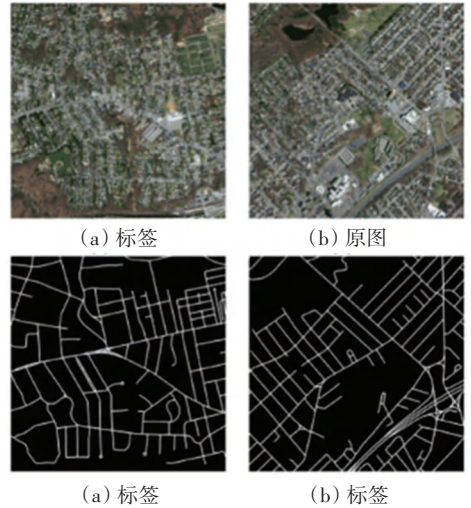


图 6 Massachusetts 道路数据集展示

Fig. 6 Massachusetts dataset presentation

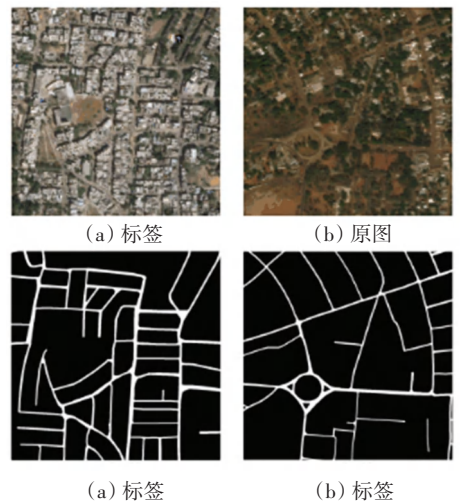


图 7 DeepGlobe 道路数据集展示

Fig. 7 DeepGlobe dataset presentation

由于训练数据集中存在影像与便签相差较大的部分, 且 GPU 内存的运算能力有限, 因此需对数据集进行预处理, 剔除信息缺失的图片。在此, 将美国马萨诸塞州道路数据集每张 1 500×1 500 像素的遥感图像切成大小为 256×256 像素图像, 将

DeepGlobe 遥感图像道路提取数据集 $1\ 024 \times 1\ 024$ 像素大小的训练图像裁剪为 256×256 像素大小的图像,切块后对两个数据集进行简单的数据增广,通过水平镜像、旋转角度、色彩抖动、模糊、增加噪音的方式进行样本扩充,数据增广效果如图 8、图 9 所示。

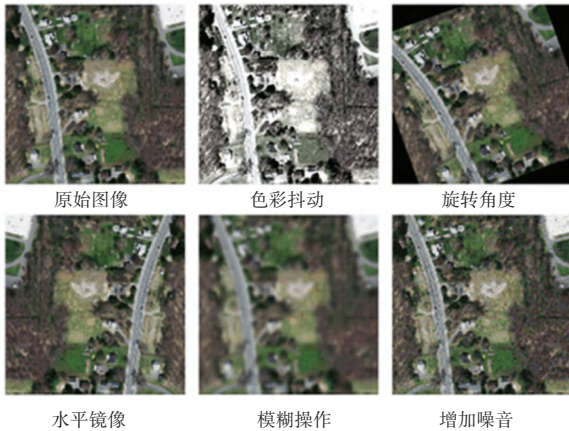


图 8 Massachusetts 数据增广效果图

Fig. 8 Data augmentation presentation

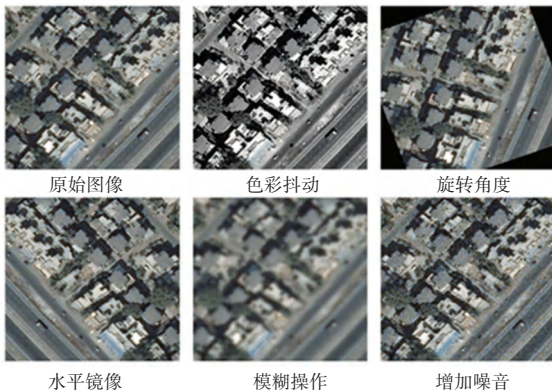


图 9 DeepGlobe 数据增广效果图

Fig. 9 Data augmentation presentation

2.3 评价指标

本文模型使用精确率 (*Precision*)、召回率 (*Recall*)、*F1 - measure* 值 (*F1* 值) 和交并比 (*IoU*) 4 项指标作为评价网络模型性能的标准,其计算公式如下:

$$Pre = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F1 = \frac{2 \times Pre \times Recall}{Pre + Recall} \quad (10)$$

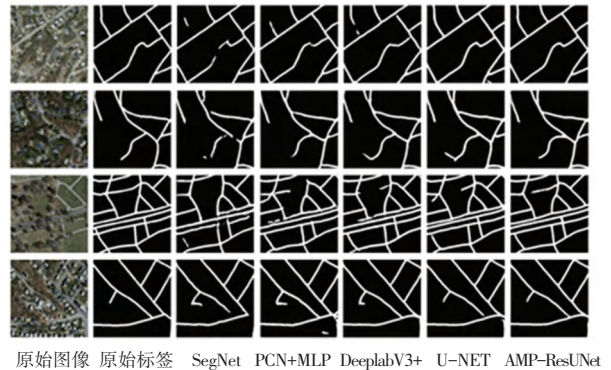
$$IoU = \frac{TP}{TP + FP + FN} \quad (11)$$

遥感图像道路信息提取,实际上是对像素进行二分类,所提取的道路信息为正样本,背景信息即为负样本。

式中: *TP* 表示实际道路被正确分类的像素数量, *FP* 表示背景像素被误分的像素数量, *TN* 表示背景像素被正确分类的像素数量, *FN* 表示实际道路像素被误分为背景的像素数量。精确率表示被正确分类的道路占总区域的比重,召回率表示被正确分类的道路占实际标注样本道路的比率, *F1* 值代表精确率与召回率之间的加权平均数,交并比与 *F1* 值可以反映预测道路信息与真实道路之间的相关性,数值越高,代表提取效果越好。

2.4 实验结果与分析

在 Massachusetts 道路数据集上对模型进行训练,本文选用经典网络模型 SegNet、FCN、DeepLabV3+、U-Net 与本文网络模型 AMP-ResUNet 做对比,实验效果如图 10 所示。



原始图像 原始标签 SegNet PCN+MLP DeeplabV3+ U-NET AMP-ResUNet

图 10 网络模型在 Massachusetts 测试集上分割效果图

Fig. 10 The network model segmented the effect picture on Massachusetts Roads Dataset

图 10 中展示的 4 幅道路图片,其背景复杂度、道路遮挡及交错情况各不相同,图中分别展示了原始图像、原始标签,以及 SegNet、FCN、DeepLabV3+、U-Net 与 AMP-ResUNet 的预测效果。从分割结果中可以看出,本文网络模型预测图较其他对比网络预测效果而言,出现错分、漏分的情况更少,对图像边缘和细节恢复的更加完整。在第三行图片中存在着树木、建筑物遮挡或边缘模糊的道路,本文网络虽也存在一定边缘不清和漏分情况,但较其他对比网络而言,本文模型对复杂道路的分割效果较好,错分、漏分情况相对较少,边缘信息恢复的更加完整,能够得到更加准确、完整的道路信息情况。

为验证改进网络在遥感图像道路分割任务上的广泛应用性,在 DeepGlobe 道路数据集上再次进行验证。该数据集中存在大量农村泥土道路和郊区道

路信息,相较于分割城市道路而言具有更高的分割难度。同上,与各类经典网络进行对比,结果如图 11 所示。

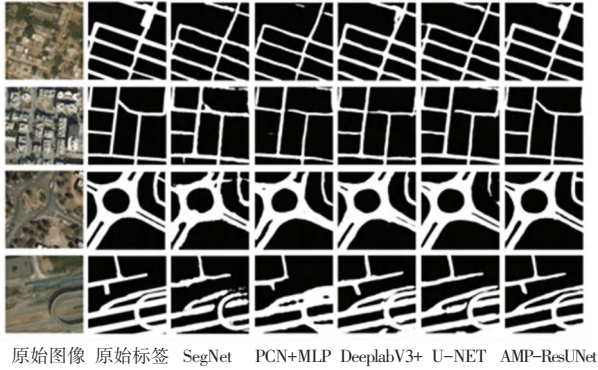


图 11 网络模型在 DeepGlobe 测试集上分割效果图

Fig. 11 The network model segmented the effect picture on DeepGlobe Roads Dataset

图 11 中展示了 4 幅图像在不同网络下的分割结果,第一行和第二行图片右上角的边缘道路和泥土道路在网络训练中较难分割,结果图片中都存在一定的漏分现象,但本文网络相较于其他对比网络而言漏分情况较少且对于道路边缘信息的恢复完整度更高。第三行和第四行图片是夹杂着泥土道路的郊区路线分布图,在 FCN 和 DeepLab 网络分割结果中可以看出,图中道路与道路间的间隔小而模糊,分割结果将两条道路混在一起,出现错分的情况。图四右上角道路被植被树木遮挡,不易分割出正确的道路,相比于其他网络而言,本文网络的分割结果良好,虽也存在边缘信息丢失的问题,但能够精准的分割出黏连的道路信息,而且对于遮挡道路也能够清晰的识别,使得道路信息恢复更加完整。

基于上文介绍的遥感图像道路提取情况的评价指标,将对比网络与本文模型网络在 Massachusetts 数据集和 DeepGlobe 数据集上的预测结果进行评价比较,具体情况见表 1、表 2。

表 1 Massachusetts 数据集上不同模型指标评价

Tab. 1 Evaluation of different model indicators on the Massachusetts Roads Dataset

Models	Pre	Recall	F1	IoU
SegNet	76.56	67.77	71.90	56.13
FCN+MLP	78.79	65.92	73.01	57.48
DeepLabV3+	79.73	67.68	73.21	57.75
U-Net	80.25	67.55	73.78	58.48
AMP-ResUNet	81.31	69.52	74.96	59.94

由表中实验数据可知,本文网络 AMP-ResUNet 在两个遥感道路数据集的双重验证下,较 SegNet、

FCN 网络在各项指标上均有大幅度提升。在 Massachusetts 数据集上,改进后的 U 型残差结构网络模型较 DeepLabV3+网络在精确率、召回率、F1 值和交并比上分别提高了 1.58%、1.84%、1.75%、2.19%。较 U-Net 网络在精确率、召回率、F1 值和交并比上分别提高了 1.06%、1.97%、1.18%、1.46%。在 DeepGlobe 数据集上,U 型残差结构网络模型较 DeepLabV3+网络在精确率、召回率、F1 值和交并比上分别提高了 3.37%、0.73%、1.40%、3.98%。较 U-Net 网络在精确率、召回率、F1 值和交并比上分别提高了 2.14%、1.34%、1.16%、3.62%。本文网络编码器结构选用 ResNet101 网络,添加了 ASPP 模块,解码器部分运用自适应混合操作以及 PixelShuffle 上采样方式,从多尺度融合深浅层次信息,在保证效果的同时提高网络整体性能。从两个数据集的实验数据中可以看出,相比于其他经典网络,AMP-ResUNet 网络训练结果的评价指标均达到最高值,充分证明了该网络模型在遥感道路分割领域上的有效性与广泛实用性。

表 2 DeepGlobe 数据集上不同模型指标评价

Tab. 2 Evaluation of different model indicators on DeepGlobe Roads Dataset

Models	Pre	Recall	F1	IoU
SegNet	86.84	76.62	81.41	68.64
FCN+MLP	89.39	71.77	79.62	66.14
DeepLabV3+	88.91	82.48	85.57	74.78
U-Net	90.14	81.87	85.81	75.14
AMP-ResUNet	92.28	83.21	86.97	78.76

3 结束语

本文对遥感图像进行道路分割研究,受残差网络、空洞金字塔池化、Adaptive Mixup 操作和 Pixelshuffle 等操作的启发,提出了一种使用超参数自适应操作调节高级特征与低级特征混合比例的残差分割网络模型 AMP-ResUNet。在编码器部分使用 ResNet101 网络保持其特征表达能力,并在一定程度上解决梯度消失或梯度爆炸问题。在编码器、解码器衔接部分引入 ASPP 模块,对特征信息进行多尺度提取。然后,在特征融合过程中 Adaptive Mixup 操作,对特征信息进行动态融合。最后,使用转置卷积与 Pixelshuffle 操作结合的方式对缩小后的特征图进行有效的放大。从预测结果图上看,本文模型对地物细节和边缘信息的提取更加完整且出现错分、漏分的情况相对较少。与其他经典语义分

割网络相比,本文网络模型在精确率、 $F1$ 值等评价指标中均达到最高值。实验表明,本文提出的 AMP-ResUNet 网络对地物信息复杂且道路遮挡严重的遥感图像有较好的分割效果,具备一定的实际应用性。在未来的工作中,将着重关注被建筑物、树木等无关信息遮挡的道路分割情况,旨在提升图像分割的准确率,实现高精度、高效率的遥感图像道路提取。

参考文献

- [1] 宝音图,刘伟,李润生,等. 遥感图像语义分割的空间增强注意力 U 型网络[J/OL]. 北京航空航天大学学报;1-11[2022-04-12].
- [2] 吴泽康,赵姗,李宏伟,等. 遥感图像语义分割空间全局上下文信息网络[J]. 浙江大学学报(工学版), 2022, 56(4): 795-802.
- [3] 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.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.
- [5] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [6] ZHANG Z, LIU Q, WANG Y. Road extraction by deep residual U-net[J]. IEEE Transactions on Geoscience and Remote Sensing Letters, 2018, 15(5): 749-753.
- [7] HE H, YANG D, WANG S, et al. Road extraction by using atrous spatial pyramid pooling integrated encoder-decoder network and structural similarity loss[J]. Remote Sensing, 2019, 11(9): 1015.
- [8] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deep Lab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [9] ZHOU L, ZHANG C, WU M. D-Link Net: Link Net with pre-trained encoder and dilated convolution for high-resolution satellite imagery road extraction[C]//In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 192-196.
- [10] CHAURASIA A, CULURCIELLO E. Link Net: Exploiting encoder representations for efficient semantic segmentation[C]//In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), 2017: 1-4.
- [11] DEMIR I, KOPERSKI K, LINDENBAUM D, et al. Deep Globe 2018: A challenge to parse the Earth through satellite images[C]//In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 172-181.
- [12] YANG X, LI X, YE Y, et al. Road detection and centerline extraction via deep recurrent convolutional neural network U-Net[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 7209-7220.
- [13] 韩彬彬,张月婷,潘宗序,等. 残差密集空间金字塔网络的城市遥感图像分割[J]. 中国图象图形学报, 2020, 25(12): 2656-2664.
- [14] 肖昌城,吴锡. 基于门控卷积残差网络的卫星图像道路提取[J]. 计算机应用研究, 2021, 38(12): 3820-3825.
- [15] 陈松钰,左强,王志芳. 基于 U-Net 的遥感图像语义分割[J]. 无线电工程, 2022, 52(1): 168-172.
- [16] WU H, QU Y, LIN S, et al. Contrastive Learning for Compact Single Image Dehazing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2021: 10551-10560.
- [17] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.
- [3] XU W, ZHOU D, HUANG X, et al. Optimal allocation of power supply systems in industrial parks considering multi-energy complementarity and demand response[J]. Applied Energy, 2020, 275: 115407.
- [4] 高岩. 智能电网实时电价社会福利最大化模型的研究[J]. 中国管理科学, 2020, 28(10): 9.
- [5] SAMADI P, MOHSENIAN-RAD A H, SCHOBER R, et al. Optimal real-time pricing algorithm based on utility maximization for smart grid[C]//2010 First IEEE international conference on smart grid communications. IEEE, 2010: 415-420.
- [6] 李军祥,张文财,高岩. 基于用户电器分类的智能电网实时定价研究[J]. 中国管理科学, 2019, 27(4): 210-216.
- [7] 张莉,高岩,朱红波,等. 考虑用电量不确定性的智能电网实时定价策略[J]. 电网技术, 2019, 43(10): 3622-3631.
- [8] 陶莉,高岩,朱红波. 以极小化峰谷差为目标的智能电网实时定价[J]. 系统工程学报, 2020, 35(3): 315-324.
- [9] ZHANG L, GAO Y, ZHU H, et al. A distributed real-time pricing strategy based on reinforcement learning approach for smart grid[J]. Expert systems with applications, 2022, 191: 116285.
- [10] LI Y, LI J, YU Z, et al. A cosh-based smoothing Newton algorithm for the real-time pricing problem in smart grid[J]. International Journal of Electrical Power & Energy Systems, 2022, 135: 107296.
- [11] 代业明. 智能电网实时电价决策的博弈论方法[D]. 上海: 上海理工大学, 2016.
- [12] 许志宏,高岩,朱红波,等. 基于分片线性效用函数的实时电价定价策略[J]. 工业工程与管理, 2018, 23(5): 44-52.
- [13] DAHLEH M, MITER S. Dynamic pricing and stabilization of supply and demand in modern electric power grids[J]. Science and Engineering, 2014, 5(12): 1130-1143.

(上接第 50 页)