李东坡，于淼，于晓鹏．CG-Res2UNet 网络模型在遥感图像道路提取中的应用[J]．智能计算机与应用，2026，16(7)：115-123．DOI：10.20169/j.issn.2095-2163.24110406

CG-Res2UNet 网络模型在遥感图像道路提取中的应用

李东坡，于　淼，于晓鹏
（吉林师范大学 数学与计算机学院，吉林 四平 136000）

摘　要：道路提取是高分辨率遥感影像处理中的一个重要课题，确保提取道路的连通性和完整性对于许多应用至关重要。然而，道路信息在复杂遥感图像中常常受到遮挡、阴影等影响，难以从单一视角有效表达全面的特征。为解决以上问题，本文提出了一种结合 CBAM 注意力机制和改进残差瓶颈结构的网络模型 CG-Res2UNet。首先，在编码器部分使用 ResNet101 深度卷积神经网络增强提取图像的多层次特征，其中融入了残差连接以缓解梯度消失问题。其次，在编码器和解码器的衔接处加入了注意力机制，使模型对特征的通道和空间维度进行自适应加权，增强了模型对关键信息的关注。最后，使用 GeLU 激活函数替代 ReLU 激活函数，引入一定程度的非线性，能够有效地捕捉输入数据中的复杂关系。为了验证所提出方法的有效性，在公开数据集 Massachusetts Roads Dataset 和 DeepGlobe Roads Dataset 上进行对比实验，与多种传统和先进的道路提取方法进行对比，结果表明本文所提的方法在多项指标上均取得了显著的提升，其中精确率分别达到了 80.23%和 76.87%。特别是在复杂场景下，如存在阴影、植被覆盖和建筑物遮挡的区域，该方法依然能够准确地提取出道路信息。

关键词：遥感图像分割；道路提取；残差网络结构；注意力机制；GeLU 激活函数
中图分类号：U491　　**文献标志码：**A　　**文章编号：**2095-2163(2026)07-0115-09

Application of CG-Res2UNet network model in road extraction from remote sensing images

LI Dongpo, YU Miao, YU Xiaopeng
(College of Mathematics and Computer, Jilin Normal University, Siping 136000, Jilin, China)

Abstract: Road extraction is an important topic in high-resolution remote sensing image processing. Ensuring the connectivity and integrity of the extracted roads is crucial for many applications. However, road information in complex remote sensing images is often affected by occlusion and shadow, which makes it difficult to effectively express comprehensive features from a single view. In order to solve the above problems, this paper proposes a network model CG-Res2UNet combining CBAM attention mechanism and improved residual bottleneck structure. Firstly, the ResNet101 deep convolutional neural network is used to enhance the multi-level features of the extracted image in the encoder part, where residual connections are incorporated to alleviate the problem of gradient disappearance. Secondly, an attention mechanism is added at the interface between the encoder and the decoder, so that the model adaptively weights the channel and spatial dimensions of the features, which enhances the model´s attention to key information. Finally, the GeLU activation function is used to replace the ReLU activation function, which introduces a certain degree of nonlinearity and could effectively capture the complex relationships in the input data. In order to verify the effectiveness of the proposed method, the paper conducts comparative experiments on the public datasets Massachusetts Roads Dataset and DeepGlobe Roads Dataset, and compares with a variety of traditional and advanced road extraction methods. The results show that the proposed method has achieved significant improvement in multiple indicators, and the accuracy reaches 80.23% and 76.87% respectively. Especially in complex scenes, such as shadow, vegetation cover and building occlusion area, this method can still accurately extract road information.

Key words: remote sensing image segmentation; road extraction; residual network structure; attention mechanism; GeLU activation function

---

基金项目：吉林省教育厅科学技术研究项目(JJKH20240570KJ)；吉林省科技发展计划项目(YDZJ202301ZYTS285)。

作者简介：李东坡(1993—)，男，硕士研究生。主要研究方向：深度学习，遥感图像分割方法。

通信作者：于　淼(1990—)，男，博士，讲师。主要研究方向：数字图像处理，深度学习。Email：ym@jlnu.edu.cn；于晓鹏(1964—)，男，教授。主要研究方向：深度学习，数值模拟研究。Email：yyxxpp@jlnu.edu.cn。

收稿日期：2024-11-04

0 引 言

道路是城乡基础设施的重要组成部分,对促进经济发展、提高居民生活质量具有重要作用。道路信息的准确提取在城市规划[1]、交通监测[2]、灾害应急[3]、环境监测[4]等各个领域都具有重要的实际应用价值。随着遥感技术的不断进步,现在已得到了更大数量的清晰图像数据[5]。高分辨率遥感图像的采集周期越来越短,这为道路的提取提供了丰富的数据集[6]。然而在高分辨率遥感图像中,道路常常会被建筑物以及道路两旁的树木所遮挡,同时由于道路上有车辆以及建筑物或树木的阴影存在也增加了道路的完整提取难度。传统人工标注的方法需要投入大量的时间和人力成本。标注者需要经过培训才能保证标注质量,而且标注过程通常比较耗时。特别是对于大规模的图像数据集,人工标注的成本可能会非常高昂[7]。因此,人们积极探索如何实现遥感图像道路的自动化提取。

深度学习(DL)技术的出现,极大地颠覆了传统的研究方法。许多基于深度学习的技术被用来提高从遥感图像中提取道路的效率,包括像 deep road Mapper[8]、拓扑损失、改善道路连通性这样的方法。其中,几个著名的卷积神经网络采用编码器-解码器架构,如 Chaurasia 等学者[9]提出的 LinkNet。利用 LinkNet 的结构和膨胀卷积,Zhou 等学者[10]引入了 D-LinkNet,该方法在 Deep-Globe 2018 道路提取挑战赛中脱颖而出。这些研究为从遥感图像中提取道路提供了重要的参考点,积极推动计算机视觉和遥感融合的进步。在语义分割这一复杂任务中,低级特征在卷积神经网络的浅层结构中得以有效捕获。传统的融合策略往往采用固定的比例来结合低级特征与高级特征,但却常常导致细节特征与语义特征出现不同程度的丢失[11]。

由此可见,仅仅依赖分割技术可能无法充分提取出道路的所有特征,特别是在地物信息复杂且道路遮挡严重的情况下。为了提高道路提取的准确性和完整性,需要综合考虑深层次语义特征和浅层纹理特征的提取融合[12]。本文受到前人对 U 形残差网络结构的研究及注意力机制[13]和特征信息融合概念的启发[14],设计出一种残差分割网络模型 CG-Res2UNet,在训练过程中自动调整超参数,能够实现浅层特征与深层特征相互融合。本文主要工作如下:

(1)以预处理过的 ResNet101 网络作为骨干网络来进行特征提取。

(2)提出了一种在单个残差块内构建分层的类残差连接块。

(3)在跳跃连接部分引入 CBAM 模块,提高网络提取重要特征的能力。

(4)采用 GeLU 激活函数以缓解梯度消失的问题。

1 CG-Res2UNet 网络工作原理

CG-Res2UNet 采用了编解码(Encoder-Decoder)结构,由多个卷积层和残差块组成(如图 1 所示)。具体而言,在编码器部分使用 ResNet101 残差网络取代原始 U-Net 网络中的编码器部分,由于残差连接的存在,网络中的信息流动更加顺畅,梯度下降算法能够更有效地更新网络权重。将原始瓶颈块中一组 3×3 卷积核替换为更小的卷积核组,以分层残差形式连接不同的卷积核组,具有更强多尺度特征提取能力的同时能够保持类似的计算负载[15]。卷积块注意力模块(Convolutional Block Attention Module,CBAM)是一种用于前馈卷积神经网络的简单且有效的注意力模块,是通道注意力(Channel Attention)和空间注意力(Spatial Attention)按顺序组合的双注意力模块[16]。在模型训练过程中,能够自动关注到不同通道和图像位置对预测结果的影响,从而使模型能够更加聚焦于关键信息,显著提升模型对细节的捕捉能力和整体的泛化性能。GeLU 激活函数通过引入高斯误差的思想,提供了一种高效且稳定的激活机制,尤其是在需要模型捕捉复杂和细微特征的任务里,为深度学习模型提供了一种有效的信息处理方式。

1.1 残差结构

随着深度学习技术的飞速发展,构建更深的神经网络成为提升模型性能的有效途径。然而,在实际训练过程中达到一定深度后,继续增加网络层次反而会导致测试错误率上升,这种现象被称为“网络退化”。残差网络(ResNet)的核心思想是通过引入残差连接(shortcut connection)来解决深度神经网络训练中的退化问题[17]。残差结构如图 2 所示,残差模块的核心思想是在网络中添加一个“跳跃连接”,使得输入可以直接跳过某些层,从而避免了在反向传播过程中梯度的衰减。同时,由于跳跃连接的存在,即使某些层的权重更新较小,也不会导致整个网络的性能下降,从而缓解了梯度爆炸问题。

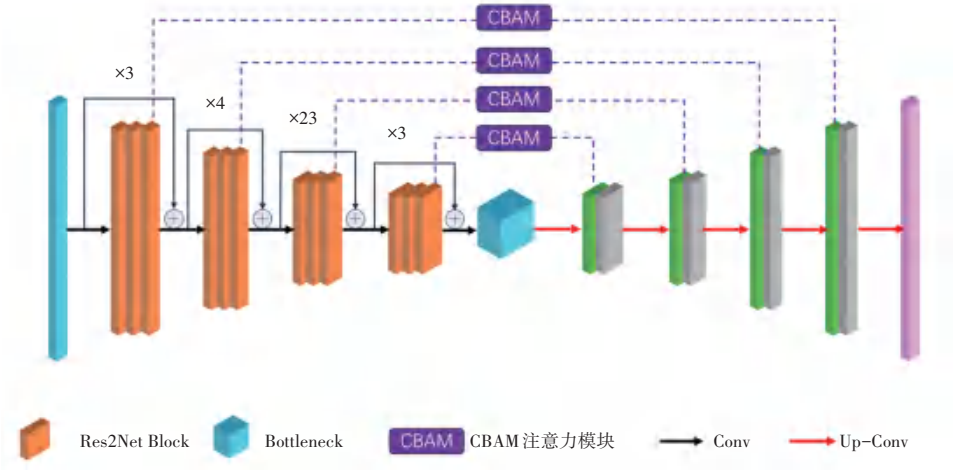


图 1　CG-Res2UNet 网络结构

Fig. 1　CG-Res2UNet network structure

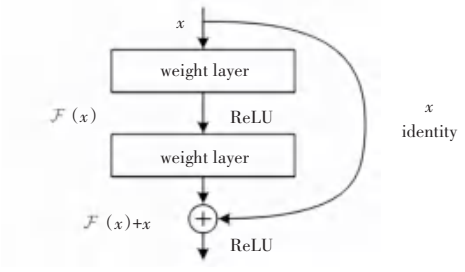


图 2　残差结构

Fig. 2　Residual structure

每个残差单元可以简洁地表达为：

$$H(x) = F(x) + x \tag{1}$$

其中，$H(x)$ 表示残差块的输出；$F(x)$ 表示输入经过残差块内部的子网络后的输出；x 表示残差块的输入。

当输入 x 与输出 $F(x)$ 维度不一致时，可表示为：

$$H(x) = F(x) + Wx \tag{2}$$

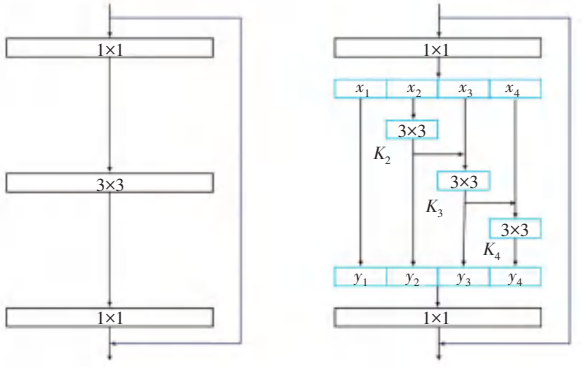
其中，W 是一个线性变换，用于调整输入 x 的维度以使其能和 $F(x)$ 进行相加操作。

1.2　改进的瓶颈结构

图 3(a)中的瓶颈结构是许多现代骨干 CNN 架构的基本构建块，例如 ResNet[18]、ResNeXt[19] 和 DLA[20]。为了寻求具有更强多尺度特征提取能力的替代架构，同时保持类似的计算负载，将一组 3×3 卷积核替换为更小的卷积核组，同时以分层残差形式连接不同的卷积核组并将其命名为 Res2Net。图 3 显示了瓶颈块和提出的 Res2Net 模块之间的差异。1×1 卷积后，将特征图均匀地分割为 s 个特征图子集，用 x_i 表示，其中 $i \in \{1,2,\cdots,s\}$。与输入特征图相比，每个特征子集 x_i 具有相同的空间大小、但通道数为 $1/s$。除 x_1 外，每个 x_i 都有对应的 3 × 3 卷积，用 $K_i(\cdot)$ 表示。同时，用 y_i 表示 $K_i(\cdot)$ 的输出。将特征子集 x_i 与 $K_{i-1}(\cdot)$ 的输出相加，然后馈入 $K_i(\cdot)$。为了在增加 s 的同时减少参数，研究中省略了对 x_1 的3 × 3 卷积。y_i 可以表示为：

$$y_i = \begin{cases} x_i, & i = 1 \\ K_i(x_i), & i = 2 \\ K_i(x_i + y_{i-1}), & 2 < i \leqslant s \end{cases} \tag{3}$$

每个 3×3 卷积运算 $K_i(\cdot)$ 都可能从所有的特征 $\{x_j, j \leqslant i\}$ 中接收到特征信息。每次特征分割 x_j 经过 3 × 3 卷积运算时，输出结果可以具有比 x_j 更大的感受野，从而在更细粒度的层面上放大感受野，增强多尺度特征学习能力。



(a) 瓶颈结构　　(b) Res2Net 模块

图 3　瓶颈块与 Res2Net 模块

Fig. 3　Bottleneck block and Res2Net module

1.3　CBAM 注意力模块

CBAM 注意力机制是一种按顺序结合通道注意力和空间注意力的轻量级模块。CBAM 注意力模块结构如图 4 所示，主要由通道注意力模块(CAM)和

空间注意力模块(SAM)组成。图 5(a)展示了通道注意力模块(CAM)的详细结构。首先,输入的特征图会分别经过全局最大池化(MaxPool)和全局平均池化(AvgPool)操作。这 2 种池化方式可以从不同角度提取特征图的信息,将尺寸为 $H \times W \times C$(这里,H 为特征图的高度,W 为宽度,C 为通道数)的特征图压缩成 2 个 $1 \times 1 \times C$ 的通道特征。然后,将其载入一个相同的多层感知机(MLP)模块,该模块由 2 个全连接层组成,中间使用 ReLU 激活函数,其中第 1 层神经元个数为 C/r,第 2 层神经元个数为 C。最后,2 个输出特征在通道维度上按对应元素相加,再应用 Sigmoid 激活函数得到输入特征图的通道注意力权重 $\boldsymbol{M}_c$,其尺寸为 $1 \times 1 \times C$。最后将原始特征与 $\boldsymbol{M}_c$ 对应的通道相乘,得到输入图像在通道维度上的标记。该模块主要用于捕捉特征图中不同通道之间的重要性差异。基于此,则会根据每个通道对最终预测结果的贡献程度,为不同的通道赋予不同的权重。

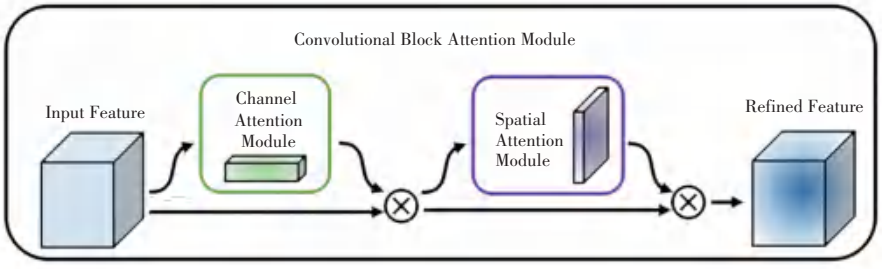
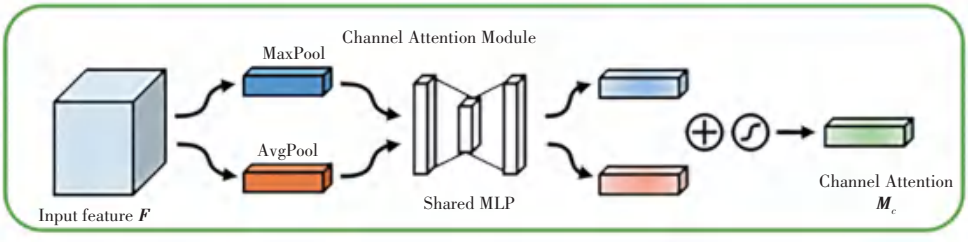
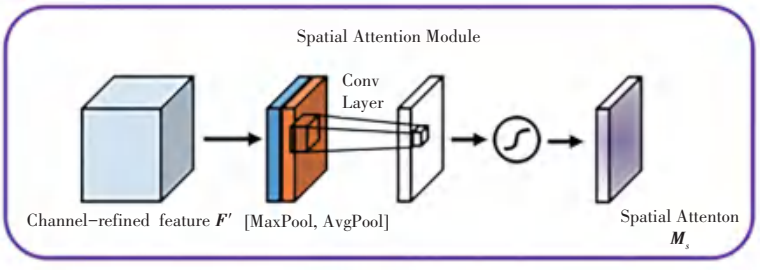


图 4 CBAM 注意力模块

Fig. 4 CBAM attention module



(a) 通道注意力模块(CAM)



(b) 空间注意力模块(SAM)

图 5 通道注意力模块和空间注意力模块

Fig. 5 Channel Attention Module and Spatial Attention Module

通道注意力权重矩阵 $\boldsymbol{M}_c$,可以表示为:$\boldsymbol{M}_c \in R^{C\times1\times1}$。为了降低计算参数,在 MLP 中采用一个降维系数 r,$\boldsymbol{M}_c \in R^{C/r\times1\times1}$。综上,通道注意力计算公式如下:

$$\begin{aligned}\boldsymbol{M}_c(F) &= \sigma(\mathrm{MLP}(\mathrm{AvgPool}(\boldsymbol{F}))) + \\ &\quad \mathrm{MLP}(\mathrm{MaxPool}(\boldsymbol{F}))) = \\ &\quad \sigma(\boldsymbol{W}_1(\boldsymbol{W}_0(\boldsymbol{F}^c_{\mathrm{avg}})) + \\ &\quad \boldsymbol{W}_1(\boldsymbol{W}_0(\boldsymbol{F}^c_{\mathrm{max}})))\end{aligned} \tag{4}$$

其中,σ 表示 Sigmoid 激活函数,$\boldsymbol{F}^c_{\mathrm{avg}}$ 和 $\boldsymbol{F}^c_{\mathrm{max}}$ 分别表示全局平均池化特征和最大池化特征。

图 5(b)展示了空间注意力模块(SAM)的结构。对经过通道注意力模块处理后的特征图,同样利用全局平均池化和全局最大池化方法,但这里是在通道维度上进行操作。全局平均池化会得到一个 $H \times W \times 1$ 的特征图,代表每个空间位置的平均特征响应;全局最大池化会得到另一个 $H \times W \times 1$ 的特征图,代表每个空间位置的最大特征响应。接下来,将

2 个特征图沿着通道维度合并，形成一个具有 $H \times W \times 2$ 维度的特征图。然后，应用尺寸为 7 × 7 的卷积核对合并后的特征图进行卷积处理，从而将其维度转换为 $H \times W \times 1$。最终应用 Sigmoid 函数对卷积后的特征图进行激活使其数值限制在 0 到 1 之间，得到空间注意力权重 $\boldsymbol{M}_s$。再将其与经过通道注意力模块处理后的特征图相乘，从而在空间维度上完成对特征的重标定[21]。在得到经过通道注意力模块处理后的特征图后，空间注意力模块会进一步关注特征图中不同空间位置的重要性差异。这一过程旨在突出特征图中对目标识别或其他任务更关键的空间区域，抑制不相关的空间区域。

空间注意力权重矩阵 $\boldsymbol{M}_s$，可以表示为：

$$\boldsymbol{M}_s(F) \in R^{H\times W} \tag{5}$$

同上，在通道维度使用 2 种池化方法产生 2D 特征图，可以表示为：

$$\boldsymbol{F}^s_{\text{avg}} \in R^{1\times H\times W} \tag{6}$$

$$\boldsymbol{F}^s_{\max} \in R^{1\times H\times W} \tag{7}$$

综上，空间注意力的计算公式如下：

$$\boldsymbol{M}_s(\boldsymbol{F}) = \sigma(f^{7\times 7}([\text{AvgPool}(\boldsymbol{F});\text{MaxPool}(\boldsymbol{F})])) = \sigma(f^{7\times 7}[\boldsymbol{F}^s_{\text{avg}};\boldsymbol{F}^s_{\max}]) \tag{8}$$

其中，σ 表示 Sigmoid 激活函数，$\boldsymbol{F}^s_{\text{avg}}$ 和 $\boldsymbol{F}^s_{\max}$ 分别表示全局平均池化特征和最大池化特征。

1.4　GeLU 激活函数

GeLU 激活函数是 Gausian Error Liner Unit 的简称，其中“高斯误差线性单元”是指一种经验模型，具有传统的线性激活函数和非线性激活函数的全部特性，可以更好地捕捉数据中的复杂模式和特征，提高模型的准确性和泛化能力。同时，GeLU 函数在一定程度上引入了随机性的因素，这有助于模型更好地应对数据中的不确定性和噪声。

GeLU 函数的原始形式定义为：

$$\text{GeLU}(x) = xP(x \leqslant x) = x\Phi(x) \tag{9}$$

其中，$\Phi(x)$ 是标准正态分布的累积分布函数，即：

$$\Phi(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \tag{10}$$

其中，erf 表示高斯误差函数。

在实际应用中，由于原始公式计算较为复杂，常使用下面近似计算公式，计算速度更快且与精确值的误差较小。该近似计算公式为：

$$\text{GeLU}(x) \approx 0.5x\left(1 + \tanh\left(\sqrt{\frac{2}{\pi}}(x + 0.044\,715x^3)\right)\right) \tag{11}$$

GeLU 激活函数如图 6 所示。

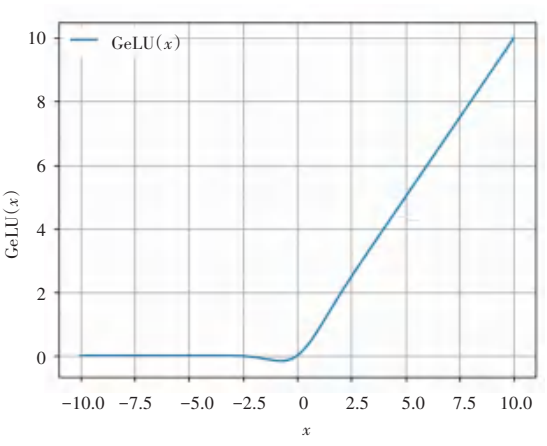


图 6　GeLU 激活函数

Fig. 6　GeLU activation function

2　实验与结果分析

2.1　实验环境

本文仿真实验代码基于 Pytorch 框架构建，编译环境为 Python3.8，操作系统为 Windows10，内存为 64 GB。硬件配置方面，CPU 型号为 Intel(R) Xeon(R) Gold5215，主频为 2.50 GHz，显卡为 NVIDIA GeForce RTX2080Ti，显存 11 GB。实验中，采用 Adam 优化策略[22]，初始学习率设为 0.000 1。为了抑制优化过程中的振荡并加速 Adam 收敛，添加了一个动量项并设置为 0.9，权重衰减用于抵抗高权重，使模型可以有更小的过拟合概率。

2.2　数据集

本研究使用马萨诸塞州遥感图像道路数据集(Massachusetts Roads Dataset)和 DeepGlobe Road 遥感图像道路数据集进行实验。其中，Massachusetts Roads Dataset 是用于道路提取和遥感影像语义分割的数据集。该数据集包含 1 171 张航拍图像，覆盖了马萨诸塞州超过 2 600 km^2 的多样化区域，包括城市、郊区和农村地区，确保数据的多样性和代表性。每张图像的尺寸为 1 500×1 500 像素，覆盖面积为 2.25 km^2。数据集被随机分为训练集(1 108 张图像)、验证集(14 张图像)和测试集(49 张图像)，数据集中的每张航拍图像都配有一张对应的二值图像，其中白色位置代表道路位置，黑色代表非道路部分。这种标注方式使得道路提取任务变得更加明确和直观，数据集的空间分辨率为 1 m/像素，即每个像素代表实际地面上的 1 m 的距离。DeepGlobe Road 数据集是一个用于道路提取的重要遥感影像数据集，包含了来自 6 个国家的高分辨率

卫星图像,覆盖了城市、乡村、荒郊、海滨、热带雨林等多种场景。该数据集的训练数据集包含 6 226 幅卫星影像,影像为 RGB 格式,大小为 1 024×1 024 像素,具有 50 cm 的像素分辨率,由 Digital Globe 的卫星收集。验证集有 1 243 幅图像,测试集有 1 101 幅图像,但测试集不提供标注信息。数据集样例如图 7 所示。

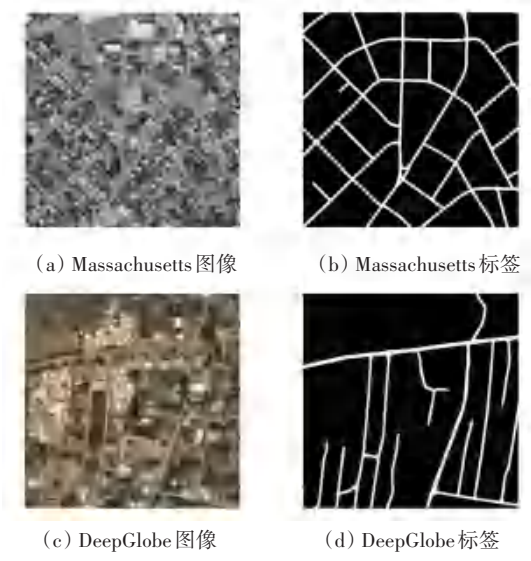


图 7 数据集图像和标签

Fig. 7 Dataset images and labels

分析可知,由于数据集的部分图像缺少与之匹配的标签,且受限于 GPU 运算能力,图像需要进行压缩处理,将图像尺寸裁剪至 512 × 512。将 DeepGlobe Road 训练集中的 6 226 张图像随机分成训练集(5 040 张图像)、验证集(560 张图像)和测试集(626 张图像)。

2.3 评价指标

为了客观地评估模型的性能,引入了 4 个评估指标,分别是精确率(Precision)、召回率(Recall)、*F*1 分数和交并比(IoU),其计算公式如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{15}$$

其中,TP(真正例)表示真实值是道路且预测值也是道路;FP(假正例)表示真实值是背景但预测为道路;FN(假反例)表示真实值是道路但预测值为背景。精确率反映了在被判断为真的样例中,实际为真的样例比例。召回率衡量的是在实际为真的样例中,被判断为真的样例比例。*F*1 分数是精确率和召回率的调和平均数,从而综合衡量了模型的性能。交并比与 *F*1 分数一起,能够揭示预测道路信息与真实道路信息之间的关联性,这两个指标的数值越高,表明道路信息的提取效果越佳。

2.4 实验结果与分析

为了评估本文提出的算法在高分辨率遥感影像道路提取任务中的有效性,研究采用了多种深度学习网络模型:FCN、LinkNet、U - Net、U - Net + +、ECAUNet,并在相同的实验条件下对 Massachusetts Roads 数据集和 DeepGlobe Road 数据集的图像进行了预测和验证。表 1、表 2 展示了经过训练得到的各模型的具体评估结果。

表 1 Massachusetts 数据集上不同模型的评估结果

Table 1 Evaluation results of different models on the Massachusetts dataset %

Models	Precision	Recall	*F*1	IoU
LinkNet	74.60	76.63	75.60	60.77
FCN	77.40	74.46	75.90	61.16
U-Net	77.98	74.29	76.09	61.41
U-Net++	78.91	73.57	76.14	61.48
ECAUNet	78.09	75.46	76.75	62.28
CG-Res2UNet	80.23	77.05	78.60	64.75

表 2 DeepGlobe 数据集上不同模型的评估结果

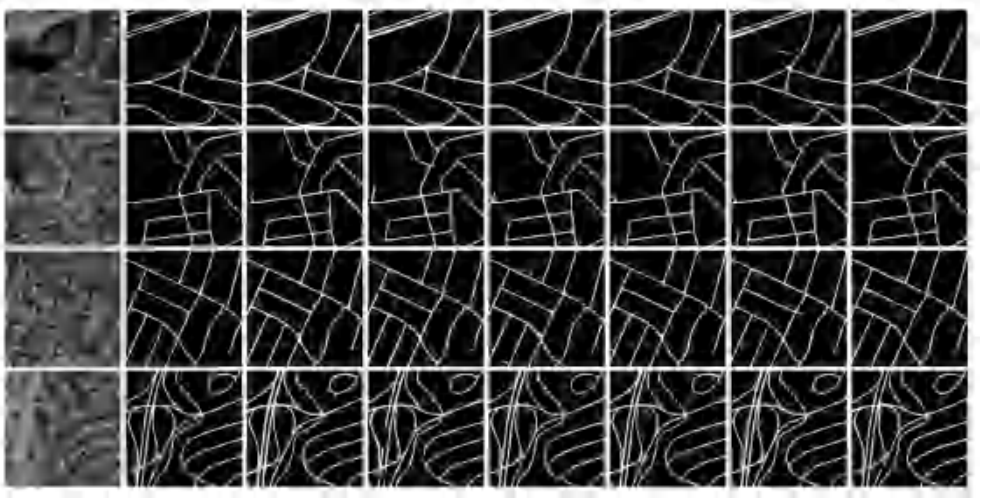
Table 2 Evaluation results of different models on the DeepGlobe dataset %

Models	Precision	Recall	*F*1	IoU
LinkNet	71.89	70.96	71.42	55.55
FCN	74.38	72.06	73.20	57.73
U-Net	72.62	74.80	73.69	58.34
U-Net++	75.63	73.99	74.80	59.74
ECAUNet	74.92	74.52	74.42	59.64
CG-Res2UNet	76.87	74.93	75.89	61.15

根据表 1 分析可知,研究提出的网络在 Massachusetts 数据集上的 Precision、Recall、*F*1 和 IoU 指标上均取得了先进结果,相较于经典网络 LinkNet、FCN 在 IoU 指标上分别提升了 3.98%、3.59%,相较于基础网络 U-Net 在 *F*1 和 IoU 指标上分别提升了 2.51%、3.34%。根据表 2 结果可以看出本文提出的网络在 DeepGlobe 数据集上的各项指标也达到了最优结果,相较于经典网络 LinkNet、FCN 在 IoU 指标上有较大提升,分别提升了 5.60%,3.42%,相较于基础网络 U-Net 在 *F*1 和 IoU 指标上分别提升了 2.20%,2.81%。实验结果证明了 CG-Res2UNet 在道路提取方面的有效性和泛化能力,比

其他对比网络有更高的识别精度。

不同网络模型在 Massachusetts 数据集上可视化结果如图 8 所示。图 8 中选取的示例图像涵盖了多种道路情况。第 1 行展示了河道旁的道路，本文网络对于河道水体附近的道路有着较高的识别度。第 2 行和第 3 行展示了城市密集道路网络，得益于 Res2Net 分层连接网络结构具有较强的多尺度特征提取能力，使网络能够以更细的粒度表达多尺度特征，从而准确地识别和提取出遥感图像中的道路特征，还原道路图像边缘细节信息。第 4 行道路图像错综复杂，本文提出加入 CBAM 注意力模块的网络模型对比其他网络模型，其提取的道路边缘平滑、且与实际道路高度吻合，减少了道路断裂、边缘粗糙等问题的出现。

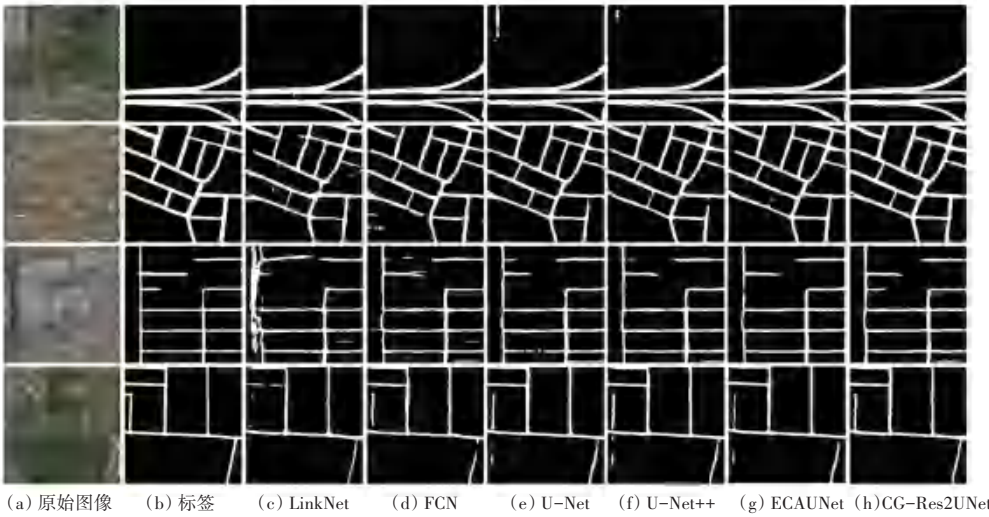


(a) 原始图像　(b) 标签　(c) LinkNet　(d) FCN　(e) U-Net　(f) U-Net++　(g) ECAUNet　(h)CG-Res2UNet

图 8　网络模型在 Massachusetts 数据集上分割结果图

Fig. 8　Graph of the segmentation results of the network model on the Massachusetts dataset

不同网络模型在 DeepGlobe 数据集上可视化结果如图 9 所示。DeepGlobe 数据集包含大量农村道路、田间泥土道路，道路与背景信息相似增加了分割难度。第 1 行左上角部分 U-Net 和 U-Net++网络存在误识别问题，本文网络能够准确区分道路与背景。第 2 行展示了农村地区房屋建筑和非铺装路面，道路与背景相似度高，可以看出本网络能在复杂背景下清晰地显示出道路细节和边缘信息。第 3 行存在河流背景信息干扰，对比网络出现不同程度将河流识别为道路的结果，本网络准确地将河流认为是背景信息并且准确识别出道路，但在右下角出现些许将房屋识别成道路的情况。第 4 行道路旁边存在大片树林，树木和树木阴影对道路遮挡严重，从对比结果观察得出本文网络分割的结果良好，整体上不存在道路中断现象，对于被遮挡的道路也能够进行清晰判断，更加准确还原出道路原始信息。



(a) 原始图像　(b) 标签　(c) LinkNet　(d) FCN　(e) U-Net　(f) U-Net++　(g) ECAUNet　(h)CG-Res2UNet

图 9　网络模型在 DeepGlobe 数据集上分割结果图

Fig. 9　Graph of the segmentation results of the network model on the DeepGlobe dataset

2.5 消融实验

为探究所提出网络模块对图像分割的有效性，分别在 2 个数据集上进行了消融实验研究。具体而言，仍然以 Precision、Recall、*F*1 值和 IoU 为评价指标，在 Baseline 基础上分别应用 GeLU 模块与 CBAM 模块，最后将两模块结合进行实验，结果见表 3、表 4。

表 3 CG-Res2UNet 在 Massachusetts 数据集上的消融研究

Table 3 Ablation study of CG-Res2UNet on the Massachusetts dataset %

Models	Precision	Recall	*F*1	IoU
Baseline	77.69	75.22	76.43	61.86
Baseline+GeLU	77.28	76.49	76.88	62.45
Baseline+CBAM	79.49	77.00	78.23	64.24
CG-Res2UNet	80.23	77.05	78.60	64.75

表 4 CG-Res2UNet 在 DeepGlobe 数据集上的消融研究

Table 4 Ablation study of CG-Res2UNet on the DeepGlobe dataset %

Models	Precision	Recall	*F*1	IoU
Baseline	73.86	73.03	73.44	58.03
Baseline+GeLU	75.02	72.59	73.79	58.46
Baseline+CBAM	76.20	74.67	75.43	60.55
CG-Res2UNet	76.87	74.93	75.89	61.15

通过表 3、表 4 中结果分析可知，使用 GeLU 激活函数对于 *F*1、IoU 指标略有提升，由于 GeLU 保留了线性激活函数并且增加了非线性激活函数，可以更好地提取数据中的复杂模式和细微特征。而 CBAM 注意力模块对于所有指标均具有较大提升，该模块通过计算每个通道的重要性权重，自适应地调整不同通道的权重，这样可以让网络更关注对当前任务重要的特征通道。同时在空间维度，该模块使模型能够在图像的不同位置上分配不同的注意力，以便在特定位置上更好地提取特征。此外通过进一步实验，将两者融合可以使模型实现更精确的分割。

3 结束语

本文研究了基于深度学习的遥感图像道路分割方法，旨在解决现有技术中存在的空间特征和地物细节信息丢失等问题。受残差网络、分层连接和注意力机制等启发，提出了基于 U-Net 网络改进的 CG-Res2UNet。通过与多种传统和先进网络分割模型相比，本文提出网络模型在精确率、回收率、*F*1 值和 IoU 等评价指标中取得显著提升。实验结果表明，本文所提方法在准确识别遥感图像中的道路方面的有效性，而且在处理不同类型的遥感图像时具有较强的鲁棒性。在未来的研究中将进一步优化算法以提高在复杂场景下的道路分割性能。

参考文献

[1] Qian D, Wang Y, Zhang X C, et al. Rationality evaluation of urban road network plan based on the EW-TOPSIS method[C]// Proceedings of International Conference on Measuring Technology and Mechatronics Automation. Piscataway, NJ: IEEE, 2021: 840-844.

[2] Shao Z D, Zheng J J, Yue G Y, et al. Road traffic assignment algorithm based on computer vision[C]//Proceedings of 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS). Piscataway, NJ: IEEE, 2023: 1-5.

[3] Huang Y S, Wei H C, Yang J L, et al. Damaged road extraction based on simulated post-disaster remote sensing images[C]// Proceedings of 2021 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ: IEEE, 2021: 4684-4687.

[4] Xu Y, Liu Shuai, Peng Y. Research and design of environmental monitoring and road lighting system based on the Internet of things [C]//Proceesings of Chinese Automation Congress. Piscataway, NJ: IEEE, 2018: 1073-1078.

[5] Cheng G, Han J W. A survey on object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 117(7): 11-28.

[6] Hormese J, Saravanan C. Automated road extraction from high resolution satellite images[J]. Procedia Technology, 2016, 24: 1460-1467.

[7] 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]. 北京航空航天大学学报, 2023, 49(7): 1828-1837.

[8] Mattyus G, Luo W, Urtasun R. DeepRoadMapper: Extracting road topology from aerial images[C]// Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2017: 3458-3466.

[9] Chaurasia A, Culurciello E. LinkNet: Exploiting encoder representations for efficient semantic segmentation [C]// Proceedings of 2017 IEEE Visual Communications and Image Processing (VCIP). Piscataway, NJ: IEEE, 2017: 1-4.

[10] Zhou L C, Zhang C, Wu M. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, NJ: IEEE, 2018: 182-186.

[11] 李梓瑜, 王大东, 于晓鹏. 基于 U 型残差网络的遥感图像道路提取方法研究[J]. 智能计算机与应用, 2023, 13(9): 51-58.

[12] 邵小美, 张春亢, 韦永昱, 等. 道路信息提取的 GF-2 影像融合方法对比分析[J]. 智能计算机与应用, 2022, 12(2): 152-157.

[13] 朱张莉, 饶元, 吴渊, 等. 注意力机制在深度学习中的研究进展[J]. 中文信息学报, 2019, 33(6): 1-11.

[14] 杜培军, 柳思聪. 融合多特征的遥感影像变化检测[J]. 遥感

学报，2012，16(4)：663-677.
[15]刘瑞明，徐春融，周韬，等. 基于深度残差网络模型压缩的局部遮挡人脸识别[J]. 智能计算机与应用，2023，13(11)：244-249.
[16]张明，张芳慧，宗佳平，等. 基于轻量级网络的人脸检测及嵌入式实现[J]. 图学学报，2022，43(2)：239-246.
[17]Shafiq M，Gu Z. Deep residual learning for image recognition：A survey[J]. Applied Sciences，2022，12(18)：8972.
[18]He K M，Zhang X Y，Ren S Q，et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2016：770-778.
[19]Xie S，Girshick R，Dollár P，et al. Aggregated Residual Transformations for Deep Neural Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2017：5987-5995.
[20]Yu F，Wang D，Shelhamer E，et al. Deep layer aggregation[J]. arXiv preprint arXiv，1707.06484，2017.
[21]侯向宁，赵金伟，黄孝斌，等. 基于双线性 RepVGG 注意力网络的花卉分类[J]. 激光杂志，2024，45(4)：165-171.
[22]马丽莹，魏云冰. 一种变分模态分解与 Adam 优化的 LSTM 电价预测方法[J]. 智能计算机与应用，2022，12(12)：142-146.