引文格式：王子明. 基于注意力机制的遥感影像道路提取方法[J]. 地理空间信息, 2025, 23(1): 13-16. doi:10. 3969/j.issn. 1672-4623.2025.01.003

基于注意力机制的遥感影像道路提取方法

王子明[1]
（1. 江苏省地质工程勘察院，江苏 南京 211102）

摘 要：针对目前深度学习算法提取遥感影像道路存在的效率和完整性较差的问题，提出了一种改进的结合空间与通道注意力机制的编码器—解码器结构。在编码器部分引入SCConv模块，从空间、通道两个方向对CNN进行压缩，以减少冗余特征，获得更高的提取效率；在编码器—解码器中间部分采用坐标注意力提高模型提取道路边界的准确性，进而提高完整性。实验表明，在马赛诸塞州数据集下改进的注意力U-Net模型的Recall、F1-score和IoU均比原始模型有明显提高，提取的道路结构更完整且耗时更短，证明了算法改进的有效性。

关键词：遥感影像；道路提取；深度学习；注意力机制
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2025）01-0013-04

Remote Sensing Image Road Extraction Method Based on Attention Mechanism

WANG Ziming[1]
(1. Geo-engineering Investigation Institute of Jiangsu Province, Nanjing 211102, China)

Abstract: Aiming at the poor extraction efficiency and completeness of current deep learning algorithms for remote sensing image road extraction, we proposed an improved encoder-decoder structure that combining space and channel attention mechanisms. We introduced the SCConv module in the encoder part to compress the CNN from space and channel directions, reducing redundant features and obtaining higher extraction efficiency. We used coordinate attention in the middle part of encoder-decoder to improve the accuracy of model in extracting the road boundary and improve the completeness. Experimental results show that the improved attention U-Net model under the Massachusetts dataset has a very obvious improvement over the original model in Recall, F1-score, and IoU. The extracted road structure is more complete and consumes less time, which proves the effectiveness of algorithmic improvement.

Key words: remote sensing image, road extraction, deep learning, attention mechanism

遥感影像分辨率的逐渐提高，使其在地物提取中发挥了不可替代的作用，尤其是针对具有不同等级特征、特征间差距较大以及细小目标较多的地物[1]。根据使用任务、功能和流量，道路可划分为高速公路、一级公路、二级公路、三级公路、四级公路5个等级；行政级别等级分为国道、省道、县道。类型繁杂且道路间光谱信息差距较大，乡村、矿区道路涵盖区域较小且与周围地物特征差异较小，难以识别提取，因此设计合理高效的提取算法至关重要。模板匹配法、形态学算法、自适应阈值分割法、知识图谱法等[2-4]传统道路提取方法往往需要人工指定规则，工作量巨大、费时费力，不能实现道路的自动化提取；而深度学习方法的自适应学习能力和特征拟合能力均十分强大，在道路提取的完整性和自动化程度上具有优势。以U-Net为代表的编码器—解码器网络在语义分割中取得了令人满意的效果，因此学者们将其应用于遥感影像道路提取中[5-8]。

1 本文方法

本文通过分析U-Net模型与注意力模块的结构特点，在U-Net模型的基础上设计了一种改进的注意力U-Net网络模型。在编码器部分引入了SCConv模块，从空间、通道两个方向对CNN进行压缩。SCConv模块包含空间重构单元（SRU）和信道重构单元（CRU）。在编码器—解码器的中间部分采用坐标注意力提高模型提取道路边界的准确性，进而提高道路信息的完整性。

1.1 U-Net模型

U-Net[9]是一种用于图像分割的深度学习模型，由Olaf Ronneberger等在2015年提出，在医学图像分割任

项目来源：山东省自然科学基金资助项目（ZR2021QD074）。
作者简介：王子明（1985—），本科，高级工程师，主要研究方向为测绘工程技术与遥感技术应用，E-mail：wuhandaxuewang2003@126.com。

务中表现出色，且被广泛应用于其他领域。U-Net模型是一种编码器—解码器网络，具有跳跃连接。编码器部分由一系列卷积层和池化层组成，用于提取图像特征，并逐渐减小空间分辨率；解码器部分由一系列卷积层和上采样层组成，用于将特征映射恢复到原始输入图像大小；跳跃连接用于连接编码器与解码器的特征映射，以帮助解码器更好地恢复细节信息。

1.2 注意力模块

通道注意力和空间注意力是注意力机制中的两种常用类型。通道注意力[10]关注特征图中不同通道之间的关系，以学习每个通道的重要性，首先计算通道维度上的平均值和最大值，再通过全连接层来学习每个通道的权重，最后将权重应用于特征图的每个通道，以增强重要的特征。空间注意力[11]关注于特征图中不同空间位置的关系，以学习每个位置的重要性，首先计算特征图上每个位置的平均值和最大值，再通过全连接层来学习每个位置的权重，最后将权重应用于特征图的每个位置，以增强重要的空间位置。

这两种类型均侧重于减少单一维度（信道或空间）上的冗余，使得网络仍存在特征冗余的问题，因此本文设计了一种新的CNN压缩方法（SCConv），以联合减少卷积层中的空间和信道冗余。SCConv由SRU和CRU组成，可有效限制特征冗余。将SCConv模块放在ResBlock[12]中3×3卷积的位置（图1）。

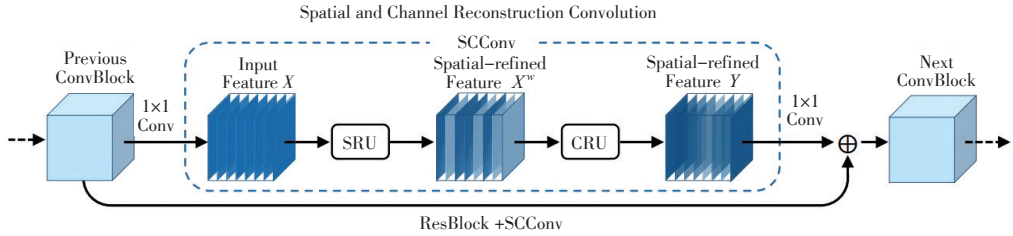


图1 SCConv在残差块中的位置

1.3 SRU部分

为减少空间冗余，本文引入了一种重构操作，首先采用交叉重构的方式充分结合加权两个信息特征，并加强它们之间的信息流；再将交叉重构的特征 X_{w1} 和 X_{w2} 连接起来，以获得空间细化的特征图 X_w（图2）。

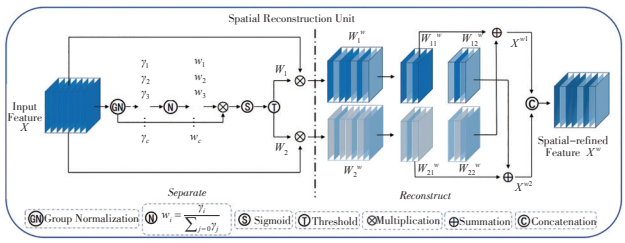


图2 空间重建单元示意图

1.4 CRU部分

CRU利用分割变换和融合策略进一步减少空间提炼特征图在通道维度上的冗余。通过Split、Transform和Fuse三个运算器实现（图3）。对于给定的空间提炼特征 $X_w \in R_{c\times h\times w}$，首先将 X_w 的通道分成为 αC 通道和 $(1-\alpha)C$ 通道两个部分，其中 $0\leqslant\alpha\leqslant1$ 为分割比；再利用 1×1 卷积来挤压特征图的通道，以提高计算效率，这里引入挤压比 r 来平衡CRU的计算成本；最后将 X_w 分成上部 X_{up} 和下部 X_{low}。

对 X_{up} 执行 $k\times k$ 分组卷积（实验中设定组大小

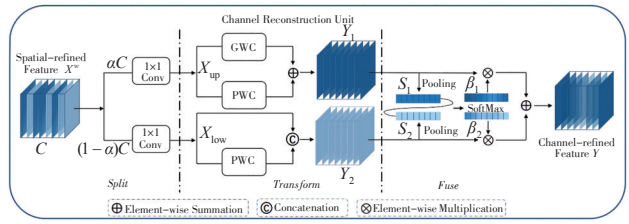


图3 信道重建单元示意图

$g=2$）和1×1逐点卷积操作，并将输出结果相加，形成一个合并的代表性特征图 Y_1，即

$$Y_1 = M^G X_{up} + M^{P_1} X_{up}$$

X_{low} 被送入较低的转换阶段，利用轻量的1×1逐点卷积操作生成具有浅层隐藏细节的特征图，作为富特征提取器的补充。此外，还可重复使用 X_{low} 特征获得更多的特征图而无需额外成本。将生成的特征和重复使用的特征串联起来，生成下级阶段 Y_2，即

$$Y_2 = M^P X_{low} \cup X_{low}$$

首先利用全局平均池化收集全局空间信息，并按信道进行统计 $S_m \in R_c \times 1 \times 1$，计算公式为：

$$S_m = \text{Pooling}(Y_m) = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} Y_c(i,j)$$

$$m = 1, 2$$

然后将上、下全局信道描述符 S_1、S_2 堆叠在一起，并采用信道软关注操作生成特征重要性向量 β_1、

$\beta_2 \in R_c$，即

$$\beta_1=\frac{e^{s_1}}{e^{s_1}+e^{s_2}},\ \beta_2=\frac{e^{s_2}}{e^{s_1}+e^{s_2}},\ \beta_1+\beta_2=1$$

最后在 β_1、β_2 的指导下，通过合并 Y_1 和 Y_2 即可得到信道提纯特征 Y，计算公式为：

$$Y=\beta_1 Y_1+\beta_2 Y_2$$

1.5　复杂性分析

SCConv模块的计算公式为：

$$P_{sc}=1\times1\times\alpha C_1\times\frac{\alpha C_1}{r}+k\times k\times\frac{\alpha C_1}{g\times r}\times\frac{C_2}{g}\times g+$$
$$1\times1\times\frac{\alpha C_1}{r}\times C_2+(1-\alpha)C_1\times\frac{(1-\alpha)C_1}{r}+$$
$$1\times1\times\frac{(1-\alpha)C_1}{r}\times\left(C_2-\frac{1-\alpha}{r}C_1\right)$$

标准卷积的计算公式为：

$$P_s=k^2C_1C_2$$

式中，α 为分割比；r 为挤压比；g 为分组卷积运算的组大小；C_1、C_2 分别为输入和输出特征通道。

在实验中通常参数集为 $\alpha=12$，$r=2$，$g=2$，$k=3$，$C_1=C_2=C$；当 $P_s/P_{sc}\approx5$ 时，参数量可减少为1/5，该模型的性能优于标准卷积。

1.6　坐标注意力模块

坐标注意力（CA）[13]模块利用池化操作 X_{pool}、Y_{pool} 分别编码水平方向和垂直方向特征，有效捕捉单方向上的长距离关系，同时保留另一个方向上的空间信息，帮助网络更准确地定位目标。特征输出的表达式为：

$$z_c^h(h)=\frac{1}{W}\sum_{0\leqslant i<W}x_c(h,i)$$
$$z_c^w(w)=\frac{1}{H}\sum_{0\leqslant j<H}x_c(j,w)$$

对水平与垂直方向特征进行拼接，经过含有正则化的卷积层后，利用通道划分使结构与经过两个一维编码的结构保持一致，再利用Sigmoid将输出激活合并为权重矩阵。每个权重都包含了通道间信息、横向空间信息和纵向空间信息，能帮助网络更准确地定位目标信息，增强识别能力。CA模块灵活且轻量，可被插入经典模块通过强化信息表示的方法增强特征；还可在轻量级网络的基础上给密集预测任务（语义分割、变化检测等）带来巨大的增益。

2　实验结果与分析

2.1　实验环境

所有实验均在PyTorch深度学习开源框架下完成，硬件配置为12 G显存的NVIDIA GeForce RTX 3060 GPU，运行环境为Python 3.8，CUDA 11.3和Torch 1.11。马萨诸塞州道路数据集是在美国马萨诸塞州收集的航空数据集，涵盖了各种典型的城市、郊区和农村地区，覆盖面积为2.25 km^2，空间分辨率为1.2 m，共有1 171幅图像，其中1 108幅用于训练、14幅用于验证、其余49幅用于测试。每幅图像大小为1 500像素×1 500像素，并带有二值分割标签。

2.2　评价指标

本文选取像素精度、F1-score和交并比（IoU）作为分类性能评估指标。

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}$$
$$F_1=\frac{2TP}{2TP+FP+FN}$$
$$\text{IoU}=\frac{TP}{TP+FP+FN}$$

式中，TP 为标签是道路、预测像素也是道路的数量；FP 为标签是背景、而预测像素是道路的数量；FN 为标签是道路、而预测像素是背景的数量；TN 为标签是背景、预测像素也是背景的数量。

2.3　结果分析

为验证改进措施的有效性，本文采用消融实验的方法，分析U-Net模型、仅加入SRU的U-Net模型、仅加入CRU的U-Net模型、加入SCConv的U-Net模型和同时加入SCConv和CA的U-Net模型在Massachusetts道路数据集上的预测效果。结果见表1和图4，可以看出，U-Net模型在解码器中融合了不同尺度的特征以保留空间信息，具有较高的总体精度，但在F1-score和IoU上略有提升空间；U-Net+CRU模型通过分割、挤压和融合策略来改进通道信息的表示，在F1-score上比U-Net模型有明显提升，在IoU上略有提高，对U-Net模型提取结果中存在的断裂现象有所改善；U-Net+SRU模型提高了特征的信息丰富性，改进了空间特征的表达，在F1-score和IoU上略有提升空间，在完整性上比U-Net+CRU有了进一步提升；U-Net+SCConv模型通过在空间维度上引入交叉信息来增强特征表达，在F1-score和IoU上都有明显提升，表明空间、信道间交叉信息有助于准确分类和定位，对于空间通道的重构、预测结果连通性与完整性至关重要；U-Net+SCConv+CA模型在U-Net+SCConv模型的基础上引入了CA，以增强位置感知和方向感知的信息，CA通过编码水平和垂直方向的信息来改善定位和识别任务，在F1-score和IoU上都有进一步提升，CA有助于更精准地定位目标信息。

表1 各网络模型道路提取结果的比较

模型	*OA*	F_1	IoU	Params/时间
U-Net	0.971 5	0.830 7	0.740 7	7.74 M/15 s
U-Net+CRU	0.975 4	0.857 5	0.773 4	7.84 M/13 s
U-Net+SRU	0.976 4	0.862 5	0.779 8	7.86 M/14 s
U-Net+SCConv	0.981 3	0.890 8	0.817 7	7.92 M/11 s
U-Net+SCConv+CA	0.981 7	0.894 3	0.822 6	9.73 M/12 s

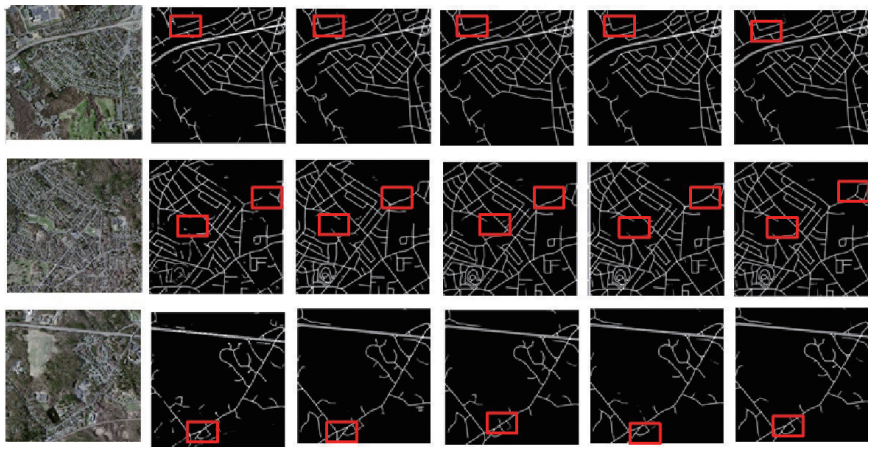


图4 模型预测结果图

3 结 语

在遥感影像道路提取领域，本文比较了U-Net、U-Net+CRU、U-Net+SRU、U-Net+SCConv和U-Net+SCConv+CA等分割算法的性能，得到的主要结论为：U-Net+CRU模型利用CRU在道路提取任务中取得了较好的性能，有效减少了道路错分和漏分现象；U-Net+SRU模型通过SRU改善了道路分割边缘的清晰度；U-Net+SCConv和U-Net+SCConv+CA模型则在模型精度和细节提取方面表现出色，特别是U-Net+SCConv+CA进一步提高了模型性能。未来道路提取领域的研究可继续探索模型精度的优化，尤其是结合道路的形态特征。此外，跨领域迁移学习、多模态信息融合等也是未来的研究方向，可进一步提高道路提取算法的适应性和实用性。

参考文献

[1] Liu Y S,Li Y H. Revitalize the World's Countryside[J]. Nature,2017,548(7 667):275

[2] 刘彦随,周扬,李玉恒.中国乡村地域系统与乡村振兴战略[J].地理学报,2019,74(12):2 511-2 528

[3] 刘殿锋,王莹莹,孔雪松,等.多维产业引导的鄂西生态圈农村居民点整理模式及分区[J].农业工程学报,2019,35(22):27-37

[4] 胡斯斌,肖建能,赵耀龙.佛冈县农村空心化判别与宅基地优化设计[J].地理空间信息,2021,19(9):143-146

[5] 王亚娟,孔福星,刘小鹏,等.中国生态移民村社会空间的生产分析:以宁夏固原市典型生态移民村为例[J].经济地理,2020,40(11):158-166

[6] Liu W,Xu J,Li J,et al. Rural Households' Poverty and Relocation and Settlement: Evidence from Western China[J]. International Journal of Environmental Research and Public Health,2019,16(14):235-244

[7] 米欢,文琦,马永霞,等.黄土高原地区贫困农户移民意愿及影响因素研究:基于原州区金轮村农户调查[J].干旱区资源与环境,2019,33(5):54-59

[8] Xu W,Jin X,Liu J,et al. Analysis of Influencing Factors of Cultivated Land Fragmentation Based on Hierarchical Linear Model: A Case Study of Jiangsu Province, China[J]. Land Use Policy,2021(101):167-174

[9] 汪冰,杨斌,李巧霞.湖北省易地扶贫搬迁贫困户地理信息系统建设[J].地理空间信息,2020,18(1):25-27

[10] 宋佳楠,金晓斌,周寅康.基于多层线性模型的耕地集约利用对粮食生产力贡献度分析:以内蒙古自治区为例[J].资源科学,2010,32(6):1 161-1 168

[11] 刘殿锋,周泊远,何建华,等.空间交互视角下中国城市群耕地变化影响因素分析[J].农业工程学报,2019,35(16):274-283

[12] 温福星.阶层线性模型的原理与应用[M].北京:中国轻工业出版社,2009

[13] 韦惠兰,罗万云.生计禀赋、环境感知与干旱区农民搬迁意愿:基于甘肃省14县市农户调查数据[J]. 干旱区资源与环境,2019,33(2):75-82