

文章编号 1004-924X(2025)10-1638-19

方向引导与拓扑感知的光学遥感道路提取网络

孟月波^{1,2*}, 黄欣羽^{1,2}, 苏世龙^{1,2}, 王 恒^{1,2}

(1. 西安建筑科技大学 信息与控制工程学院, 陕西 西安 710055;

2. 建筑机器人陕西省高等学校重点实验室, 陕西 西安 710055)

摘要: 针对光学遥感影像道路提取过程中连通性弱、细微分支遗漏以及预测路网与真实路网拓扑结构不一致等问题, 本文提出了一种方向引导与拓扑感知的道路提取网络。首先, 设计多路径方向引导模块关注多方向连接关系, 分离并独立学习不同方向的连通特征, 增强道路分支间的联系, 提高分割连续性; 其次通过全粒度互补特征融合模块融合细、粗粒度特征的互补信息, 强化道路细节与语义表达, 提高网络对细微分支的捕捉能力; 最后设计了拓扑感知损失函数, 旨在从拓扑视角探索几何结构差异, 增强模型对拓扑变化的敏感度, 约束预测路网与真实路网的拓扑一致性。所提模型在 DeepGlobe 与 Massachusetts 数据集上的 F1 值分别为 81.95% 和 79.98%, 对比现有最佳方法提升了 0.73% 和 1.5%, IoU 为 69.35% 和 66.38%, 提升了 0.98% 和 0.66%。实验结果表明, RDTA-Net 提取出的道路精确度和完整度优于其他主流方法。此外, 在遮挡、噪声以及光照等复杂场景下, 仍具有优越的提取效果。

关键词: 道路提取; 连通关系; 方向引导; 全粒度特征融合; 拓扑感知

中图分类号: TP394.1; TH691.9 **文献标识码:** A

doi: 10.37188/OPE.20253310.1638

CSTR: 32169.14.OPE.20253310.1638

Optical remote sensing road extraction network with directional guidance and topological awareness

MENG Yuebo^{1,2*}, HUANG Xinyu^{1,2}, SU Shilong^{1,2}, WANG Heng^{1,2}

(1. College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;

2. Key Laboratory of Construction Robots for Higher Education in Shaanxi Province, Xi'an 710055, China)

* Corresponding author, E-mail: mengyuebo@163.com

Abstract: To address the challenges of weak connectivity, subtle branch omission, and topological inconsistency between predicted and real road networks in optical remote sensing image road extraction, this paper proposed a road extraction network with directional guidance and topological awareness. First, the multi-path directional guidance module was designed to model multi-directional connectivity relationships. By decoupling and independently learning connectivity features across distinct directions, this module enhanced inter-branch linkages and improved segmentation continuity. Second, the full granular complementary feature guidance module integrated fine-grained and coarse-grained features, reinforcing both road de-

收稿日期: 2024-11-10; 修订日期: 2025-01-12.

基金项目: 国家自然科学基金面上项目 (No. 52278125)

tails and semantic representations to strengthen the network's capability in capturing subtle branches. Finally, a topological awareness function was introduced to quantify geometric structural discrepancies from a topological perspective, thereby constraining the topological consistency between predicted and real road networks. The proposed model achieves F1 scores of 81.95% and 79.98% on the DeepGlobe and Massachusetts datasets, outperforming the state-of-the-art methods by 0.73% and 1.5%, respectively. The IoU metrics reach 69.35% and 66.38%, with improvements of 0.98% and 0.66% over existing benchmarks. Experimental results demonstrate that RDTA-Net significantly surpasses mainstream methods in both road extraction accuracy and completeness. Furthermore, it exhibits robust performance in complex scenarios involving occlusions, noise, and illumination variations.

Key words: road extraction; connectivity; directional guidance; full granular feature fusion; topology awareness

1 引言

随着遥感技术日益发展,光学遥感影像数据的获取能力与成像质量不断提升,在海量的数据支持下,道路提取技术得到蓬勃发展,该技术利用遥感影像提取出道路网络信息,在自动驾驶、环境监测、灾后救援^[1]和路网规划等领域具有重要意义,为城市智能交通和地理信息系统建设提供可靠的数据基础。通过分析近十几年以来提出的遥感道路提取算法,可分为基于传统算法的道路提取方法与基于深度学习的道路提取方法。

在遥感道路提取技术的初期,研究者侧重于利用手工特征来实现道路的提取。Shao 等人^[2]根据像素的灰度值与带状特征的位置关系,将各像素划分为道路类和背景类。Laptev 等人^[3]提出了一种基于道路的多尺度检测方法,利用流形几何约束边缘来构建路网结构。Wegner 等人^[4]设计了一种用于道路提取的新型 CRF 模型,其中先验由沿直线段连接超像素集的高阶派来表示,一定程度提升了提取精度。上述传统道路提取方法提取效果较差,同时在面临遮挡时效率低下,难以满足任务需求。

随着深度学习的快速发展,基于卷积神经网络的模型^[5-8]被广泛地应用在道路提取技术中。其中研究者利用 U-Net^[7]与 LinkNet^[8]出色的上下文信息获取能力,提出了一系列针对道路特性的改进变体。Xu 等人^[9]提出了一种基于 U-Net 的道路提取方法,并引入局部和全局注意单元来

识别不同宽度的道路。Zhou 等人^[10]利用空洞卷积扩大感受野,加强对道路特征的理解,提高分割精度。这些方法增强模型在复杂环境下对道路特征的获取,但缺乏对道路连通性的关注,导致分割路网断裂。为此,Yang 等人^[11]通过位置注意机制与上下文信息提取策略,提高对遮挡路段的推理能力。Kampffmeyer 等人^[12]首次对像素关系进行建模,增强预测结果的整体连通性。Mei^[13]等人探究像素间依赖关系来感知道路形状,改善道路连续性不佳的情况。上述文献大多采用逐像素分类,将每个像素视为独立个体,从本质上忽视了道路的连通关系;同时,这些方法缺乏对方向特征的捕捉,导致不同方向道路分支难以建立联系,进而影响预测路网的连通性。因此,准确捕捉道路间的连通关系及其方向信息显得尤为关键。

此外,为解决道路受遮挡且与其他地物特征相似,导致的细微道路分支丢失问题,研究者们致力于为模型提供更详细的特征表示,提高分割结果完整性。Jie 等人^[14]采用多尺度特征编码与长程上下文感知模块,提高模型对细长道路的识别能力。吴等人^[15]通过坐标卷积和全局信息增强模块,减轻空间特征和上下文信息丢失,确保道路分类的精确性。Hou 等人^[16]提出了一种补充 Unet 的道路提取方法,引入多尺度密集扩张卷积获取更精细的道路特征,缓解细小道路漏分情况。Pang 等人^[17]设计语义增强策略来优化浅层特征表达,并利用特征注意策略来捕捉细节信息。然而,上述方法仅依赖于单一粒度的特征图

来补充道路细节,且在融合阶段采用特征拼接或逐像素相加的方式,未能考虑不同粒度特征图之间的表征差异,限制网络对道路特征的整体理解,导致细微分支遗漏问题。

受地理条件的影响,光学遥感道路具有复杂的拓扑形态,常用的像素级损失缺乏拓扑约束力,导致网络难以正确推理道路几何形态。Mo-sinska 等人^[18]设计了一种损失函数获取曲线结构的描述,加强对拓扑形态的理解。Shit 等人^[19]通过计算分割掩码及其形态骨骼交叉部分的相似性,增强模型对结构变化的关注。为进一步增强网络对曲线结构感知,Hu 等人^[20]与 Wong 等人^[21]利用拓扑学理论来反映结构的全局关系,来考量拓扑视角相似度,大幅提高预测分割的拓扑正确性。上述文献局限性在于仅关注形状规则的目标,难以适应道路长跨度、不规则的特性;此外,仅关注各节点与其最近节点间的拓扑差异,未考虑所有拓扑特征点间的关联,导致模型只能修正拓扑差异较小的形态,对于显著拓扑错误缺乏约束,致使预测路网与真实路网间存在较大偏差。

针对上述挑战,本文提出了方向引导与拓扑感知的道路提取网络(Road Extraction Network

with Directional Guidance and Topological Awareness, RDTA-Net)。首先,设计了多路径方向引导模块(Multi-Path Directional Guidance Module, MDG),旨在探索道路像素间的连通关系,并将不同方向信息进行分离后独立学习,加强方向特征的表达,提高路网的整体连通性;其次,在跳跃连接处引入全粒度互补特征融合模块(Full Granular Complementary Feature Fusion Module, FGC),深挖不同粒度特征间的互补信息,优化道路细节与语义融合,有效减少细微道路分支的遗漏;最后,提出一种拓扑感知损失函数(Topological Awareness Loss Function, TALoss),旨在利用拓扑学理论衡量预测路网与真实路网的形态差异,增强模型对拓扑错误的敏感性,从而提升光学遥感图像中道路结构的推理能力。

2 方 法

在本节中,本文设计了一种用于从光学遥感图像中提取道路的网络(RDTA-Net),其整体结构如图 1 所示,包括五个部分:特征编码器,MDG,FGC,TALoss以及特征解码器。

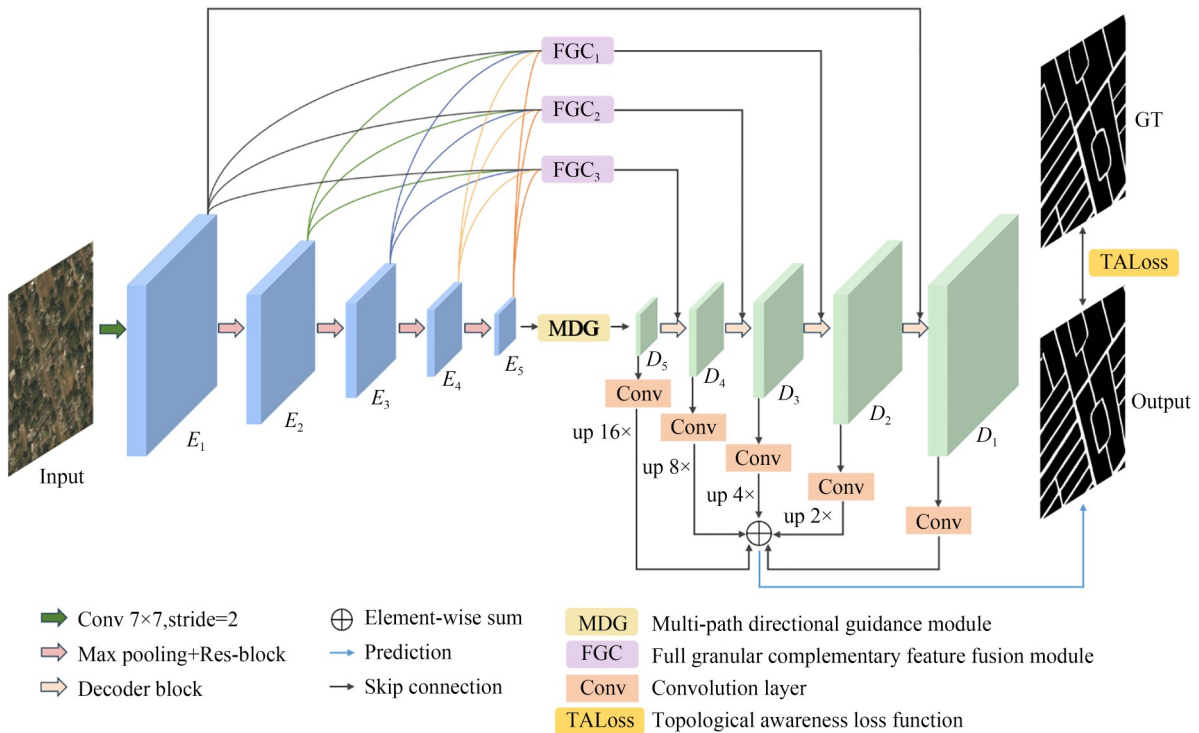


图 1 RDTA-Net 总体结构图

Fig. 1 Overall structure of RDTA-Net

2.1 特征编码器

语义分割模型中编码器是关键组件,其语义与位置信息的捕捉能力直接影响提取效果。ResNet 凭借深层残差学习结构展现出优异的特征提取能力,成为众多主流算法的编码器基础;考虑到 ResNet-34 适中计算复杂度,更适用于光学遥感地物提取任务。因此,本文采用 ResNet34 作为编码器来提取多层次的特征信息,并保留了所有阶段提取的特征图,为解码阶段提供可靠的语义和位置信息的支持。编码器的输出表示为 E_m , m 代表编码器第 m 层, $m \in (1, 2, 3, 4, 5)$ 。首先将原始遥感图像通过核大小为 7×7 、步幅为 2 的卷积得到第一层编码输出 E_1 ,随后多次通过 Maxpooling 与残差块来提取道路特征。

2.2 多路径方向引导模块

由于道路具有特殊的连通结构,以及不同区

域的道路形态存在较大差异,大多网络采用逐像素分类的方式对路网进行分割,割裂了像素间的连接关系,且缺乏对道路方向特征的感知,导致模型难以捕获道路的连通结构,造成分割路网连通性较弱。本文探索像素的连通关系^[12],显式地加强了模型对道路连通结构的关注,以解决逐像素分类方式对道路连通性的弱关注问题;此外,本文进一步研究发现,连接掩码通道中蕴含了丰富的方向特征,与光学遥感道路的曲线结构相契合。因此,本文设计了如图 3 所示的 MDG 将通道各方向特征进行有效分离,并为不同的方向构建特定的学习路径,为曲线结构的提取开辟全新视角,让不同方向的道路分支建立联系,从而提高道路网络的整体连续性。具体过程如表 2 所示。

首先,本文通过如表 1 所示的步骤,推理各像素 (x, y) 与其相八方向像素 $C_1 \sim C_8$ 的连通关系,得到连通向量 V_c ,并将其表示为如图 2 所示的连

表 1 道路连通关系的推理

Tab. 1 Reasoning about road connectivity

算法 1:道路连通关系的推理步骤

Input: road_mask($H \times W$)	#传统分割掩码
Output: connectivity_mask $G(H \times W \times 8)$	#8 方向连通掩码
1: $G \leftarrow \text{ZeroTensor}(H, W, 8)$	#初始化 G , 形状为 $(H, W, 8)$, 所有元素设为 0
2: $d \leftarrow [(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0), (1, 1)]$	#定义 8 个方向的相对偏移 (dx, dy)
3: $x \leftarrow 0$	
4: while $x < H$	
5: $y \leftarrow 0$	
6: while $y < W$	
7: if road_mask[x, y] == 1	#仅处理道路像素
8: $d = 0$	
9: while $d < 8$	
10: $dx, dy \leftarrow \text{directions}[d]$	#获取方向偏移
11: $nx = x + dx$	
12: $ny = y + dy$	#计算邻居坐标
13: if $0 \leq nx < H$ and $0 \leq ny < W$ and road_mask[nx, ny] == 1	#检查邻居是否在图像边界内且为道路
14: $G[x, y, d] = 1$	#方向 d 处连通
15: end	
16: $d \leftarrow d + 1$	#处理下一个方向
17: end	
18: $y \leftarrow y + 1$	#处理下一列
19: end	
20: $x \leftarrow x + 1$	#处理下一行
21: end	
22: Return G	#返回方向连通掩码

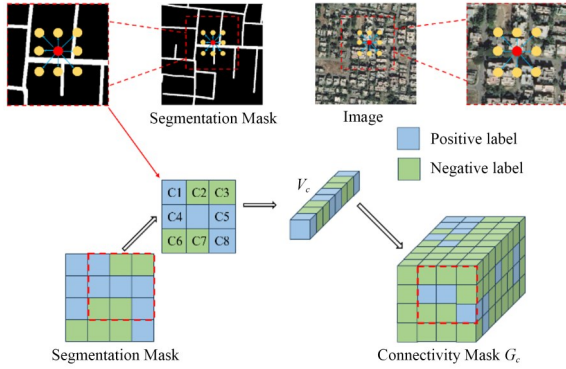


图 2 连接 mask 推理过程

Fig. 2 Connection mask reasoning process

接掩码 $G_c \in R^{H \times W \times 8}$, G_c 各通道中不仅表示各像素所属类别,还反映了连接方向。例如,在 i 方向上,若两个像素均属于道路,即 $G_{ci}(x, y) = 1$ 。因此,相较于分割掩码,连接掩码具有更丰富的方向感知,且随网络的深入, G_c 蕴含的方向特征自然以通道的形式储存。

其次,为进一步挖掘通道间方向特征,本文利用全局平均池化操作对 E_5 进行压缩,并使用 1×1 的卷积 W_l 进行降维,得到 $v_{con} \in R^{E_5}$;之后对 v_{con} 采用 1×1 的卷积 W_2 与 Sigmoid 激活函数,进行再编码与归一化,得到方向向量 d_{con} ,包含了道路各方向的重要程度。计算过程如下:

表 2 多路径方向引导算法

Tab. 2 Multi-path directional guidance algorithm

算法 2:多路径方向引导算法

Input: 特征图 E_5 , (i, j) 表示各像素位置, H_e, W_e 代表 E_5 的高度与宽度,

Output: 方向感知特征图 E_{MDG}

Step 1: $d_{con} \leftarrow \sigma \left(W_2 \delta \left(W_l \frac{1}{H_e \times W_e} \sum_{i=1}^{H_e} \sum_{j=1}^{W_e} E_5(i, j) \right) \right)$ # 计算多种方向之间的依赖关系,得到方向向量 d_{con} , W_l 表示卷积,

δ 与 σ 分别表示 ReLu 与 Sigmoid 激活函数

Step 2: 按照道路八向关系将 d_{con} 与 E_5 进行通道切片,得到 $d_{con}^j, E_5^j, j \in (1, 2, \dots, 8)$

Step 3: For j to 8 # 为各方向构建特定学习路径,给予各方向不同的关注度

Step 3.1: $E_5^j \leftarrow \text{Conv}(\text{CPAM}(E_5^j) + \text{CCAM}(E_5^j))$ # 捕获空间、通道维度依赖关系

Step 3.2: $E_{con}^j \leftarrow W_3^j F_c(d_{con}^j, E_5^j) + E_5^j$ # 根据 d_{con}^j 中该方向的重要性,引导网络给予各方向不同的关注度, F_c 表示通道乘法, W_3^j 表示卷积

Step 4: $E_{MDG} \leftarrow \text{Conv}(\text{concat}(E_{con}^1, E_{con}^2, \dots, E_{con}^8))$ # 将各方向输出 E_{con}^j 按通道进行拼接,得到方向感知特征图 E_{MDG} , 算法结束

$$\text{GAP}(E_5) = \frac{1}{H_e \times W_e} \sum_{i=1}^{H_e} \sum_{j=1}^{W_e} E_5(i, j), \quad (1)$$

$$v_{con} = \delta(W_l \text{GAP}(E_5)) = \delta \left(W_l \frac{1}{H_e \times W_e} \sum_{i=1}^{H_e} \sum_{j=1}^{W_e} E_5(i, j) \right), \quad (2)$$

$$d_{con} = \sigma(W_2 v_{con}), \quad (3)$$

其中: H_e 和 W_e 分别是 E_5 的高度与宽度, δ 表示 ReLu 激活函数, $\text{GAP}()$ 为全局平均池化操作, $W_l \in R^{C_{E_5}/16 \times C_{E_5}}$, $W_2 \in R^{C_{E_5} \times C_{E_5}/16}$, σ 代表 Sigmoid 激活函数。

接着,为了有效分离道路不同方向特征,按八向连接关系对 d_{con} 与 E_5 进行通道切片,得到

$d_{con}^j, E_5^j, j \in (1, 2, \dots, 8)$, 同方向的切片构成一个路径, E_5^j 分别通过紧凑性注意力模块^[22]捕获空间、通道维度的上下文依赖关系,将结果进行串联融合后使用卷积操作来降维,得到 E_5^j 。该步骤通过为各方向信息构建特定的学习路径,加强道路方向特征的表达。

然后,对 E_5^j 与 d_{con}^j 进行通道乘法给予各方向不同的关注度,使用 1×1 卷积 W_3^j 进行再编码后,将 E_5^j 直接添加到输出上得到 E_{con}^j 。其计算过程如下:

$$E_5^j = \text{Conv}(\text{CPAM}(E_5^j) + \text{CCAM}(E_5^j)), \quad (4)$$

$$E_{con}^j = W_3^j F_c(d_{con}^j, E_5^j) + E_5^j, \quad (5)$$

其中: $\text{Conv}()$ 代表卷积操作, $\text{CPAM}()$ 表示紧

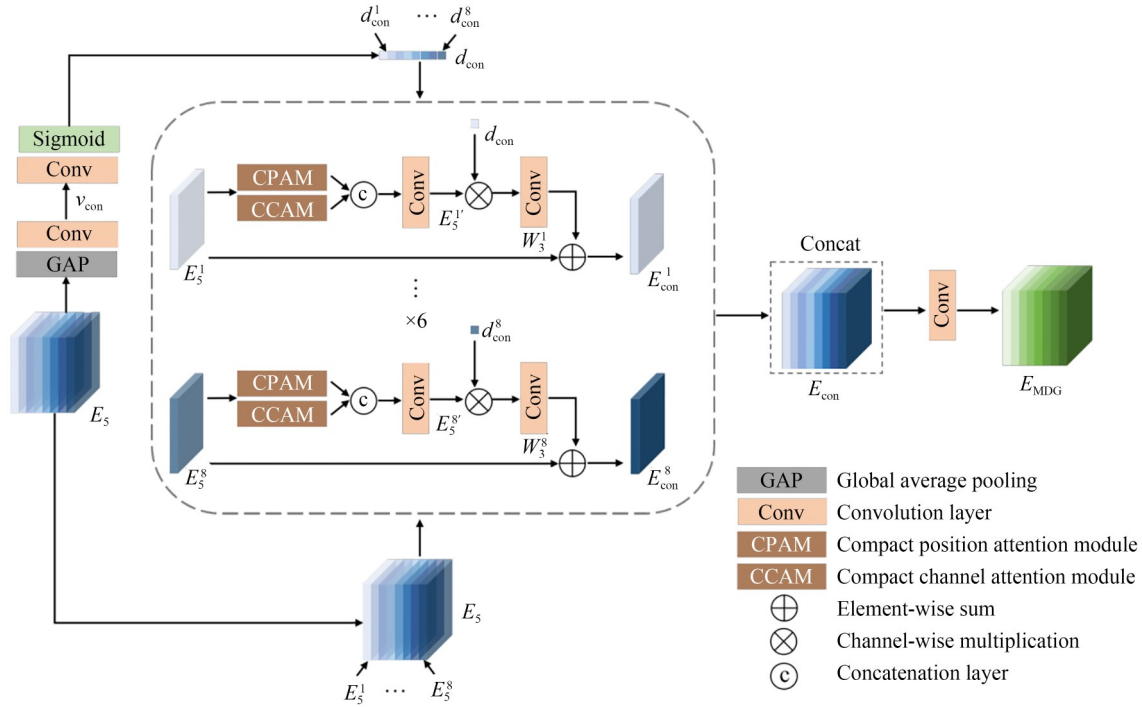


图 3 多路径方向引导模块

Fig. 3 Multi-Path Directional Guidance Module

湊性空间注意力模块,CCAM()表示紧凑型通道注意力模块, F_c 表示特征切片 E_5^i 与方向切片 d_{con}^i 之间的通道乘法。

最后,按通道维度对八条路径的方向输出 E_{con}^i 进行拼接,并使用卷积操作得到方向感知特征图 E_{MDG} ,为后续解码阶段提供丰富的方向特征与连通感知,计算过程如下:

$$E_{MDG} = \text{Conv}(\text{concat}(E_{con}^1, E_{con}^2, \dots, E_{con}^8)), \quad (6)$$

其中,concat()表示通道拼接操作。

2.3 全粒度互补特征融合模块

由于网络在编解码特征融合阶段只依赖单粒度特征来补充道路细节,对多粒度特征融合不充分,使得网络无法全面理解细小道路的特征,导致细微分支遗漏问题。因此,本文在跳跃连接处嵌入了如图4所示的FGC,旨在剖析不同粒度特征图中的多尺度互补信息,加强细粒度道路细节与粗粒度道路语义的融合,获得对道路特征的全表达以解决环境干扰造成的细微分支遗漏问题。

以FGC₁例,首先,考虑到语义粒度差距的影响^[17],将 E_m 分为细粒度特征图 E_1 与 E_2 、粗粒度特征图 E_4 和 E_5 与主特征图 E_3 ;通过卷积操作将所有特征图的尺寸与通道数调整至与 E_3 一致;其

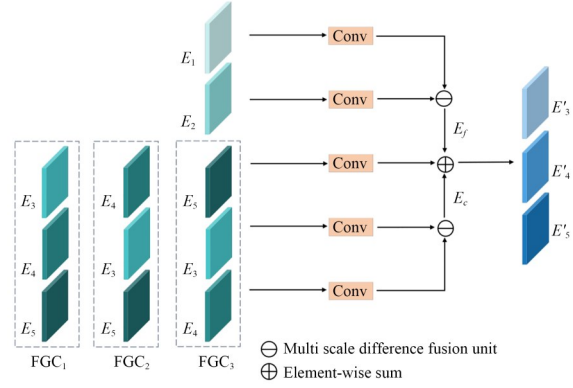


图 4 全粒度互补特征融合模块

Fig. 4 Full granular complementary feature fusion module

次,通过多尺度差异融合单元(Multi Scale Difference Fusion Unit, F_{mdu}),分别计算 E_1, E_2 以及 E_4, E_5 间的多尺度互补信息,用于发掘道路细节与语义特征,具体计算过程如下所示:

$$E_f = F_{mdu}(E_1, E_2), \quad (7)$$

$$E_c = F_{mdu}(E_4, E_5), \quad (8)$$

$$F_{mdu} = \text{Conv}(|W_{1 \times 1}(E_A) \ominus W_{1 \times 1}(E_B)| + |W_{3 \times 3}(E_A) \ominus W_{3 \times 3}(E_B)| + |W_{5 \times 5}(E_A) \ominus W_{5 \times 5}(E_B)|), \quad (9)$$

其中: E_f, E_c 分别表示道路细节与语义, E_A 与 E_B 为特征图, $Conv(\cdot)$ 表示卷积层, $\ominus$ 是逐像素减法运算, $|\cdot|$ 计算绝对值, $W_{n \times n}$ 表示尺寸为 $n \times n$ 的卷积。 F_{mdu} 通过 $W_{n \times n}, n \in (1, 3, 5)$ 三组尺寸不同的卷积,来捕获道路不同尺度的结构互补特征,并将它们进行融合,从而获得道路特征的全面表达。

最后,本文利用逐元素加法的操作来融合 E_3, E_f 与 E_c ,融合后的特征图 E'_3 将参与后续解码,提供丰富的全粒度多尺度互补信息,以实现道路特征的精准理解。计算过程如下:

$$E'_3 = E_3 + E_f + E_c. \quad (10)$$

2.4 拓扑感知损失函数

在光学遥感道路结构中,几何与拓扑信息是推断路网结构的关键,如何实现对形态的约束在于从拓扑角度反映道路结构的差异。因此,本文设计了一种基于持续同调理论^[23](Persistent Ho-

mology, PH)的拓扑感知损失函数,旨在追踪道路拓扑特征的演变,为模型补充拓扑视角信息,并综合考量所有特征点的拓扑偏差,有效捕捉并惩罚大幅度的形态差异,增强模型对拓扑错误的敏感性。

具体过程如图5所示。将原始光学遥感影像输入至RDTA-Net得到分割结果,利用PH追踪尺度参数 ϵ 变化期间拓扑特征的演变,来计算预测与真实路网的特征点,并将其表示在持久图(Persistent Diagram, PD)中,该图反映了不同维度的道路拓扑结构;接着,衡量两个持久图的拓扑差异,并深挖所有异常点的拓扑线索,进一步增强模型对几何结构变化的敏感性;最后将拓扑损失与像素级损失联合使用,从全局拓扑约束与局部细节精准性两个层面对网络进行双重监督,实现对道路形态的正确推理。如表3所示为TALoss的计算过程。

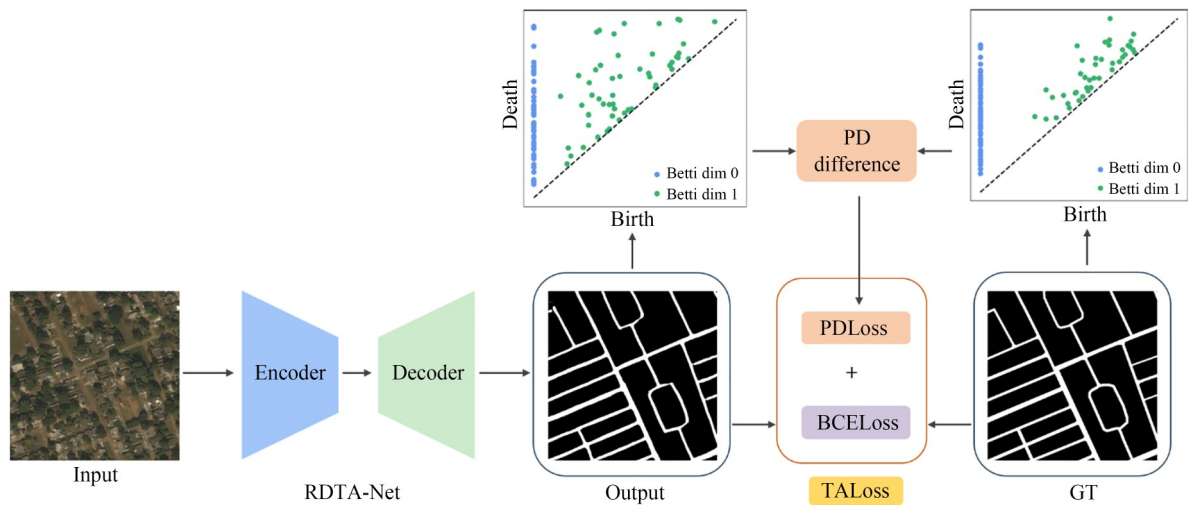


图5 拓扑感知损失函数示意图

Fig. 5 Schematic diagram of topology-aware loss function

具体来说,给定 X ,其 K 维拓扑结构称为同调类^[24,25],是 K 维流形的等价类,这些流形可以在 X 中相互转换,维度 $k \in 0, 1, 2$ 。特别地,0维和1维持续同调特征分别代表连接组件与循环,路网断裂与拓扑不一致的线索蕴含在这两个维度中。因此,本文的方法使用0维与1维持续同调特征来推理路网拓扑。如图6(a)所示,首先,统计尺度参数 ϵ 变化期间的道路拓扑变化,用于构建多尺度拓扑结构;接着,记录不同维度各拓扑特征

的诞生时间 b 、死亡时间 d 以及特征存活时间^[21],并表示为如图6(b)所示的持续条形码(0维和1维同源类的拓扑特征分别用蓝色与绿色条表示,彩图见期刊电子版);然后,将持续条形码转换为如图6(c)所示的持久图PD,用于描述道路的拓扑结构。其中每个点 (b, d) 表示第 k 个持久同调类在 b 时间诞生并在 d 时间死亡, $PD = Dgm(*)$ 表示为从Ground Truth(G)或输出(O)中得到的持久图。

表 3 拓扑感知损失函数计算流程

Tab. 3 Calculation process of TALoss

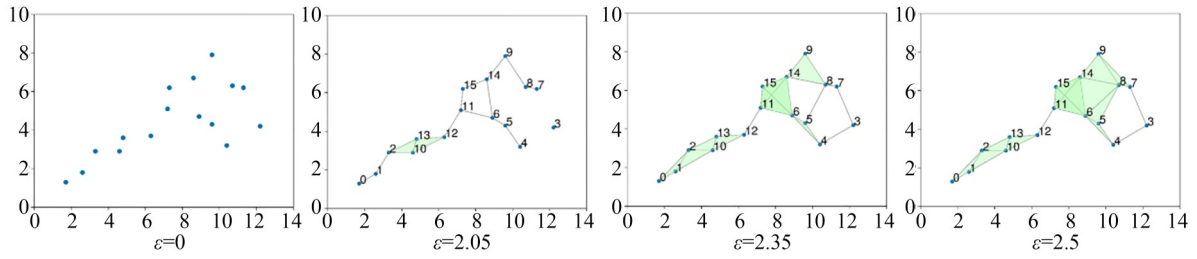
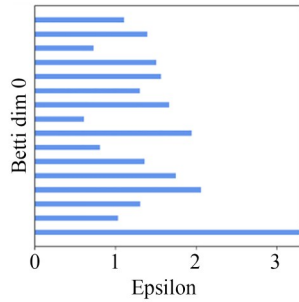
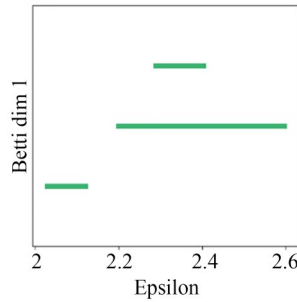
算法 2: 拓扑感知损失函数步骤**Input:** 预测结果 O , 标签 G **Output:** TALossStep 1: 计算尺度参数 ϵ 变化期间的复形序列 O_N, G_N Step 1.1: 寻找 O 与 G 中不同维度的拓扑特征点 $(\epsilon_b^i, \epsilon_d^i)$, 并表示为持续条形码 $PH_i^k(G), PH_i^k(O)$ $\# \epsilon_b^i, \epsilon_d^i$ 分别代表第 i 个特征点的出生时间与死亡时间, $k \in 0, 1, 2$, 表示特征维度Step 1.2: 利用 $PH_i^k(G), PH_i^k(O)$, 计算持久图 PD_O^k, PD_G^k Step 1.3: $L_{PD}^k \leftarrow \text{Hausdorff}(PD_O^k, PD_G^k)$ $\#$ 利用平均 Hausdorff 距离计算预测路网与真实路网的拓扑相似度Step 2: $L_{PD} \leftarrow L_{PD}^0 + L_{PD}^1$ $\#$ 统计两个维度的拓扑差异Step 3: $\text{TALoss} \leftarrow L_{PD} + L_{\text{BCE}}$ $\#$ 联合像素级损失与拓扑损失, 得到 TALoss(a) 改变 ϵ 时拓扑特征的演变
(a) Evolution of topological features as ϵ changes(b) 0维与1维持续条形码
(b) 0-dimensional and 1-dimensional persistent barcodes(c) 持久图
(c) Persistence diagram

图 6 持续同调的例证

Fig. 6 Examples of persistent homology

然后, 本文利用平均 Hausdorff 距离^[26] 综合计算 $Dgm(G)$ 与 $Dgm(O)$ 两个持久图中, 所有持久点对之间的拓扑相似度, 以更全面的视角来感知几何形态偏差, 并避免了因某异常值过大导致较大惩罚的问题, 具备更好的稳健性。定义为:

$$d^k(PD_O^k, PD_G^k) = \frac{1}{M} \sum_{m \in PD_O^k} \min_{n \in PD_G^k} \|m - n\|, \quad (11)$$

$$d^k(PD_G^k, PD_O^k) = \frac{1}{N} \sum_{n \in PD_G^k} \min_{m \in PD_O^k} \|n - m\|, \quad (12)$$

$$L_{PD}^k(PD_O^k, PD_G^k) =$$

$$\max \{d^k(PD_O^k, PD_G^k), d^k(PD_G^k, PD_O^k)\}, \quad (13)$$

其中: $d^k(PD_O^k, PD_G^k)$ 与 $d^k(PD_G^k, PD_O^k)$ 表示 k 维的有向平均豪斯多夫距离, $k \in (0, 1)$, $PD_O^k \in Dgm(O)$, $PD_G^k \in Dgm(G)$, M, N 分别代表 PD_O^k, PD_G^k 中的持久点数, m, n 分别代表 PD_O^k, PD_G^k 中的点, $L_{PD}^k(PD_O^k, PD_G^k)$ 表示在 k 维的平均豪斯多夫距离。如以上公式所示, 若预测持久图 PD_O^k 与真实持久图 PD_G^k 拓扑相似性较高, 则它们

中大多数点重合, PD_O^k 内的拓扑异常点远离 PD_G^k 中的所有点, 这些异常点反映着拓扑错误的程度, 以 $L_{PD}^k(PD_O^k, PD_G^k)$ 的大小表示。

因此, 本文对 0 维与 1 维进行求和得到 L_{PD} , 旨在重点探索这两个维度的拓扑变化, 增强对拓扑错误的敏感度。具体过程如下:

$$L_{PD} = L_{PD}^0 + L_{PD}^1. \quad (14)$$

最终的拓扑感知损失 L_{TA} , 通过联合二元交叉熵损失 L_{BCE} 与拓扑损失 L_{PD} , 使网络在拓扑特性以及像素准确性方面得到双重约束, 有效保持分割路网的拓扑正确性。具体计算过程如下:

$$L_{TA} = L_{BCE} + L_{PD}. \quad (15)$$

2.5 特征解码器

解码器的输出表示为 D_m , m 代表编码器第 m 层, $m \in (1, 2, 3, 4, 5)$ 。解码器的输入由 FGC 模块的输出和前一个解码器输出构成。前一阶段的特征图通过上采样和卷积调整其尺寸与通道数, 通过卷积、批量归一化和修正线性单元的处理, 并与全粒度道路互补信息进行融合, 生成该解码器阶段的特征图。

在预测中, 本文将特征图 D_5, D_4, D_3, D_2 通过上采样和卷积操作调整至与 D_1 同等的分辨率与通道数, 将这些特征图融合后进行上采样操作以及 3×3 的卷积以获得最终的预测连接输出。

3 实验与结果分析

3.1 数据集

本文使用 DeepGlobe 与 Massachusetts 遥感道路数据集来验证所提出的框架的性能。DeepGlobe 数据集收集自泰国、印度尼西亚和印度等国家的 8 570 张遥感影像, 每张遥感影像的尺寸为 $1\,024 \times 1\,024$, 其空间分辨率为 0.5 m/pixel , 共涵盖 $2\,220 \text{ km}^2$ 土地面积。遵循以往研究^[13], 本文使用 6 226 张标记图像作为实验数据集, 并将其划分为 4 696 张图像的训练集和 1 530 张图像的测试集, 并将每个图像裁剪为 512×512 的尺寸。Massachusetts 数据集包含 1 108 张训练图像、14 张验证图像和 49 张测试图像, 其中每张图像尺寸为 $1\,500 \times 1\,500$ 且空间分辨率为 1 m/pixel , 共涵盖了 $2\,600 \text{ km}^2$ 土地面积。本文通过填充原始影像的右侧和下侧, 得到尺寸为 $1\,536 \times 1\,536$ 的影

像, 然后将每个图像裁剪为 512×512 的尺寸。

3.2 评估指标

对于道路提取任务常用二分类评价指标来评估模型的性能, 例如准确率 (Precision)、召回率 (Recall)、 F_1 值 (F1-score) 与交互比 (IoU)。将道路定义为正样本, 背景定义为负样本。上述指标通过混淆矩阵进行计算, 其中 TP, FN, FP, TN 的含义如表 4 所示:

表 4 混淆矩阵

Tab. 4 Confusion matrix

样本	预测为道路	预测为背景
真实道路	TP	FN
真实背景	FP	TN

Precision 表示预测道路像素中真实道路的比例。Recall 表示真实道路像素被正确预测的比例。F1-score 是一种综合性指标, 被定义为准确率与召回率的调和平均数。IoU 表示道路像素的预测结果与真实标签间的相似程度。上述指标公式如下所示:

$$Precision = \frac{TP}{TP + FP}, \quad (16)$$

$$Recall = \frac{TP}{TP + FN}, \quad (17)$$

$$F_1\text{-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2 \times TP}{2 \times TP + FN + FP}, \quad (18)$$

$$IoU = \frac{TP}{TP + FP + FN}. \quad (19)$$

此外, 为了客观验证 TALoss 对道路拓扑的约束性, 本文遵循文献^[19]来计算基于拓扑的分数, 包括 Betti 数 β_0 与 β_1 的绝对 Betti 误差的平均值, 揭示预测路网与真实路网的拓扑差异, β_0 与 β_1 的值越小, 代表预测与真实路网拓扑一致性越高。

3.3 实验设置

本文所有实验均使用 NVIDIA GeForce RTX 2080 Ti 下进行, 深度学习框架采用 Pytorch 1.10.2 完成模型构建、训练与测试。在网络训练过程中, 设置训练参数 batch_size 为 8, epoch 设置为 200 轮, 初始学习率为 5×10^{-4} , 使用 AdamW 优化器进行网络优化, 并采用余

弦退火的学习率衰减策略。为提高模型的泛化能力,本文使用随机翻转、水平翻转、随机旋转、对比度调整以及高斯模糊等数据增强方法。

3.4 实验结果分析

3.4.1 分析 DeepGlobe 道路数据集实验结果

在 DeepGlobe 数据集上,本文对比方法包括 FCN^[5], SegNet^[6], U-Net^[7], D-LinkNet^[10], PSPNet^[27], RADANet^[28], CoANet^[13], SDU-Net^[29], BDTNet^[30], CR-HR-RoadNet^[31], SASwin Transformer^[32], CFRNet^[33]。

各方法的定性评估结果如图 7 所示。绿框指示断裂与道路漏检区域,红框指示分割效果提升区域。结果表明, RD TA-Net 在路网提取方面优于所有对比方法。在第一、二行红框中,由于道路与田垄等地物特征相似,其他方法易遗漏细微分支,而 RD TA-Net 通过 FGC 模块获得更丰富的道路特征表达,实现了更完整的提取效果。第三行中, FCN, SegNet 等方法因未建模道路连通性,导致明显断裂;引入 MDG 模块后,模型对道路方向更加敏感,有效建立了分支间联系。第四行中,在复杂几何结构下,其他方法出现严重拓扑错误,而 TALoss 增强了模型对拓扑结构的感

知能力,成功恢复真实路网形态。第五行中,路网中的辅路模糊易混淆,对比方法出现断裂和错分, RD TA-Net 凭借对光学遥感中细节的良好感知能力,有效提取出了辅路。然而,在第三行和第五行中,道路位于高密度城区之中,且其纹理特征与周围地面或田垄高度相仿,导致本文所提网络未能识别绿框区域中的交叉路段,出现漏检现象。该结果表明,尽管网络在多数场景下具有良好的提取能力,但在类田地背景干扰较强的复杂区域仍存在一定局限。后续可结合更精细的光谱区分特征或引入辅助信息进一步提升模型的判别能力。

对各方法进行定量分析结果如表 5 所示。本方法在 Precision, Recall, F1-score 和 IoU 指标分别为 80.69%, 83.26%, 81.95% 和 69.35%。相较于其他先进方法,本文提出的方法在召回率、F1-score 和 IoU 获得最高的分数。 BDTNet 通过注意力与特征细化策略来保持路网完整性,使准确率获得最高的分数。 RD TA-Net 更专注于道路的连续性与拓扑结构,因此其召回率远远优于 BDTNet,表明本文的方法对于总体路网的漏检情况较少。 RD TA-Net 的 F1-score 与 IoU 分别优于最佳比较方法 0.73% 与 0.98%。

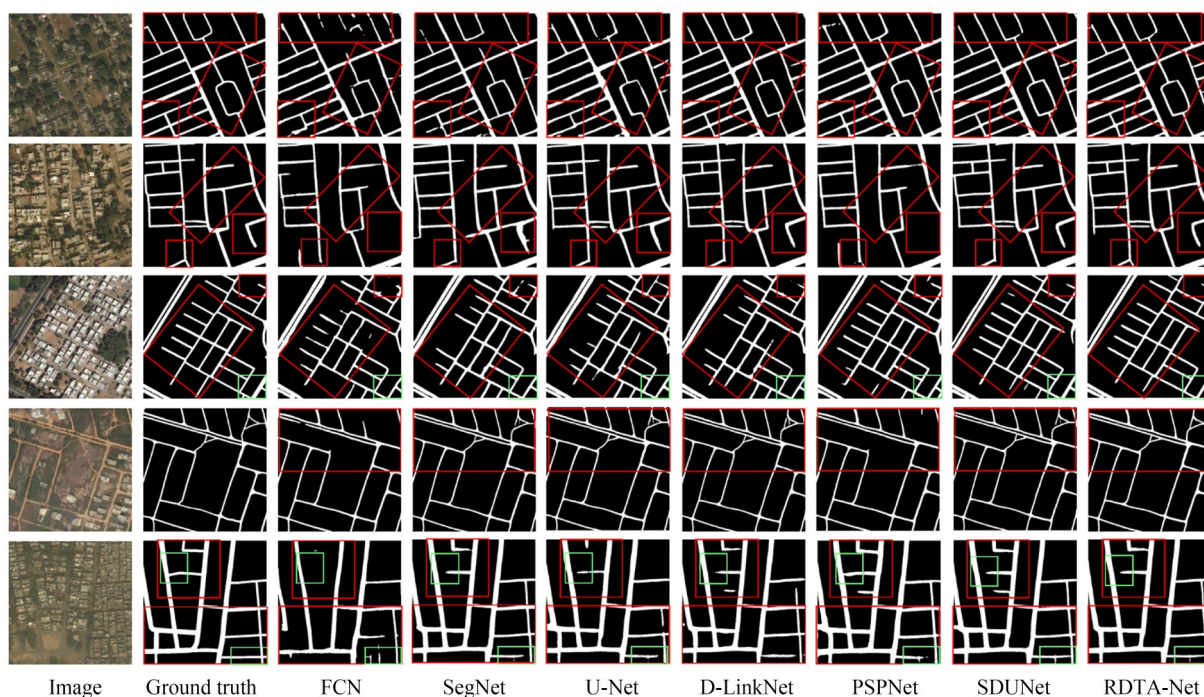


图 7 DeepGlobe 数据集各种模型的道路提取结果

Fig. 7 Road extraction results of various models in DeepGlobe dataset

表 5 DeepGlobe 数据集的定量评估结果

Tab. 5 Quantitative evaluation results on the DeepGlobe dataset (%)

Methods	Precision	Recall	F1-score	IoU
FCN	70.53	73.65	72.06	56.32
SegNet	69.60	78.17	73.64	58.28
U-Net	72.70	77.66	75.10	60.12
PSPNet	66.27	79.48	72.28	56.47
D-LinkNet	74.64	80.03	77.24	64.55
SDUNet	78.40	74.20	79.40	66.80
RADANet	/	/	73.67	/
CoANet	/	/	81.22	68.37
BDTNet	84.18	76.77	80.30	67.09
CR-HR-RoadNet	76.47	77.12	76.79	62.33
SASwin Transformer	80.97	77.56	79.23	65.60
CFRNet	81.03	79.77	80.40	67.22
RDTA-Net	80.69	83.26	81.95	69.35

3. 4. 2 分析 Massachusetts 道路数据集实验结果

Massachusetts 数据集上,本文的对比方法包括 FCN^[5], SegNet^[6], U-Net^[7], D-LinkNet^[10], PSPNet^[27], SDUNet^[29], CR-HR-RoadNet^[31], SASwin Transformer^[32], CADUNet^[34], SGCN^[35], DDU-Net^[36], CRAENet^[37], D-FusionNet^[38]。

各方法定性评估结果如图 8 所示(彩图见期刊电子版)。绿框指示断裂与道路漏检区域,而红框指示分割效果提升的区域。第一行道路受植被遮挡而辨识度低,其他模型对道路特征理解不佳,导致分割路网出现道路漏分、误判的情况, RDTA-Net 通过 FGC 模块将遥感图像中辨识度较低的道路特征解析出来,在复杂环境干扰下仍能提取出细微分支;第二、第三行中对比方法难以感知道路的形态结构。TALoss 通过比较预测路网与真实路网的拓扑差异,从而推理出正确的道路形态,但在绿框区域中的路段结构复杂不规则,本文方法未能完整提取出路段。这一结果表明 TALoss 在多数场景下能显著提升道路形态的推理能力,但在面对环境干扰强烈、结构多样的区域时仍存在感知盲区,后续可结合高分辨率特征增强机制或多源信息融合策略进一步提升网络的鲁棒性。第四行红框中,不同区域的道路具有多样的连接方式,对比模型在感知道路连通结构时存在局限性,造成预测路网连通性不佳,所提模型专注于探究道路间多向连通关系,从而提取出连续性路网;第五行道路被严重遮挡,对比方法难以有效提取道路特征, RDTA-Net 深挖全粒度互补信息获取更精细的道路特征,分割路网

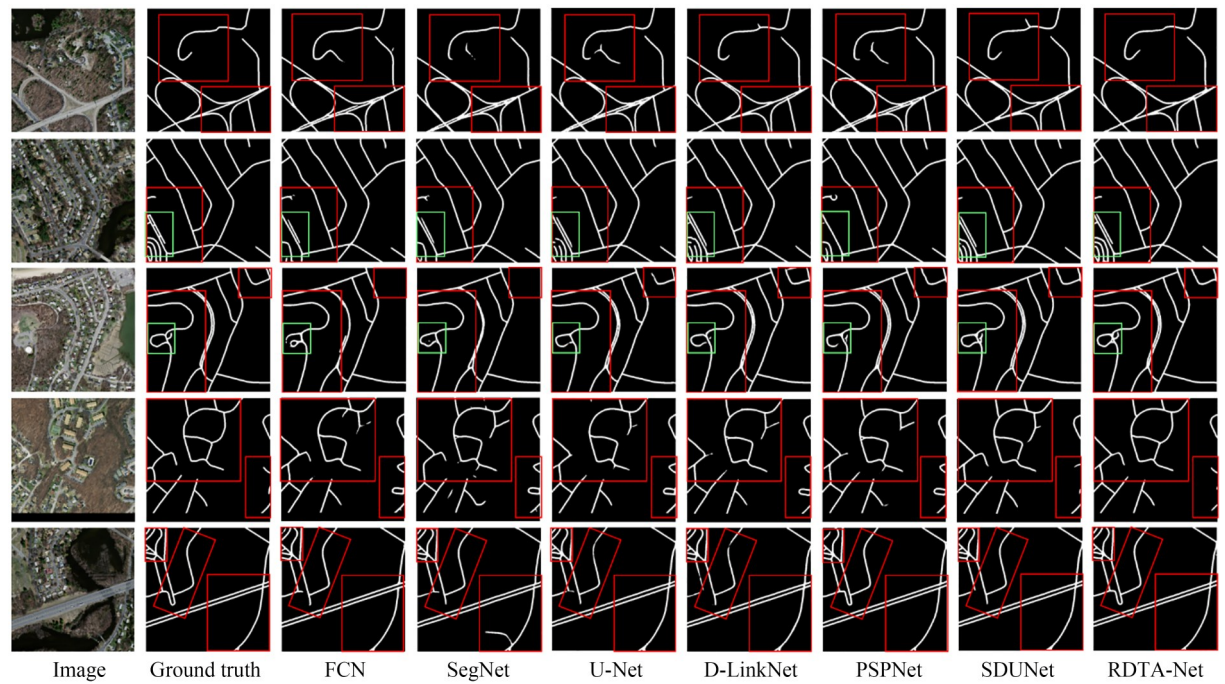


图 8 Massachusetts 数据集各模型的道路提取结果

Fig. 8 Road extraction results of each model in Massachusetts dataset

的完整度优于其他方法。

定量评价结果如表 6 所示。RDTA-Net 在 Precision, Recall, F1-score 与 IoU 上分别达到 79.91%, 80.05%, 79.98% 和 66.38%, 并在召回率、F1 值和 IoU 上获得最佳分数。特别在 Recall 指标上优于最佳比较模型 2.57%, 表明 RDTA-

表 6 Massachusetts 数据集的定量评估结果

Tab. 6 Quantitative evaluation results on the Massachusetts dataset (%)

Methods	Precision	Recall	F1-score	IoU
FCN	68.52	68.94	68.73	52.16
SegNet	74.33	77.14	75.71	60.32
U-Net	75.81	78.73	77.24	62.54
PSPNet	75.92	77.33	76.61	61.49
D-LinkNet	76.99	78.22	77.60	63.45
CR-HR-RoadNet	80.34	76.15	78.19	64.19
SASwin Transformer	80.02	77.65	78.82	65.04
CADUNet	79.45	76.55	77.89	64.12
SGCN	84.82	73.91	78.99	65.28
DDU-Net	82.54	73.99	78.03	63.98
CRAENet	81.21	77.48	78.48	65.72
D-FusionNet	/	/	78.11	64.08
RDTA-Net	79.91	80.05	79.98	66.38

Net 对曲折的道路结构具有更好的感知力, 分割路网连续性更优异。F1 值和 IoU 指标上分别优于最佳方法 1.5% 与 0.66%。

3.4.3 消融实验与分析

为验证所提模型与模块的有效性, 在 DeepGlobe 与 Massachusetts 数据集上开展消融实验。Baseline 是带有预训练 ResNet34 的基本 U 形模型, L_{TA} 为本文的 TALoss。Model1 为 Baseline, Model2 表示基于 Baseline 添加 MDG; Model3 表示基于 Baseline 添加 FGC; Model4 表示基于 Baseline 同时添加 MDG 和 FGC; Model5 表示基于 Model4 添加 L_{TA} 。

如图 9 所示, 前两行和后两行分别展示了 DeepGlobe 与 Massachusetts 数据集上消融实验的可视化结果。在第一行红框内 Model1 的分割路网严重的道路断裂、细微分支遗漏及拓扑错误问题。Model2 红框内, MDG 的添加有效引导模型对道路连接关系的关注, 显著提高预测路网的连通性; FGC 的引入进一步提升了分割完整性, 成功提取出部分被遮挡路段; Model4 同时采用 MDG 与 FGC, 充分探索道路连通结构和特征表达, 大幅提高道路整体提取效果; 在此基础上 Model5 通过引入 TALoss, 使网络聚焦于道路的

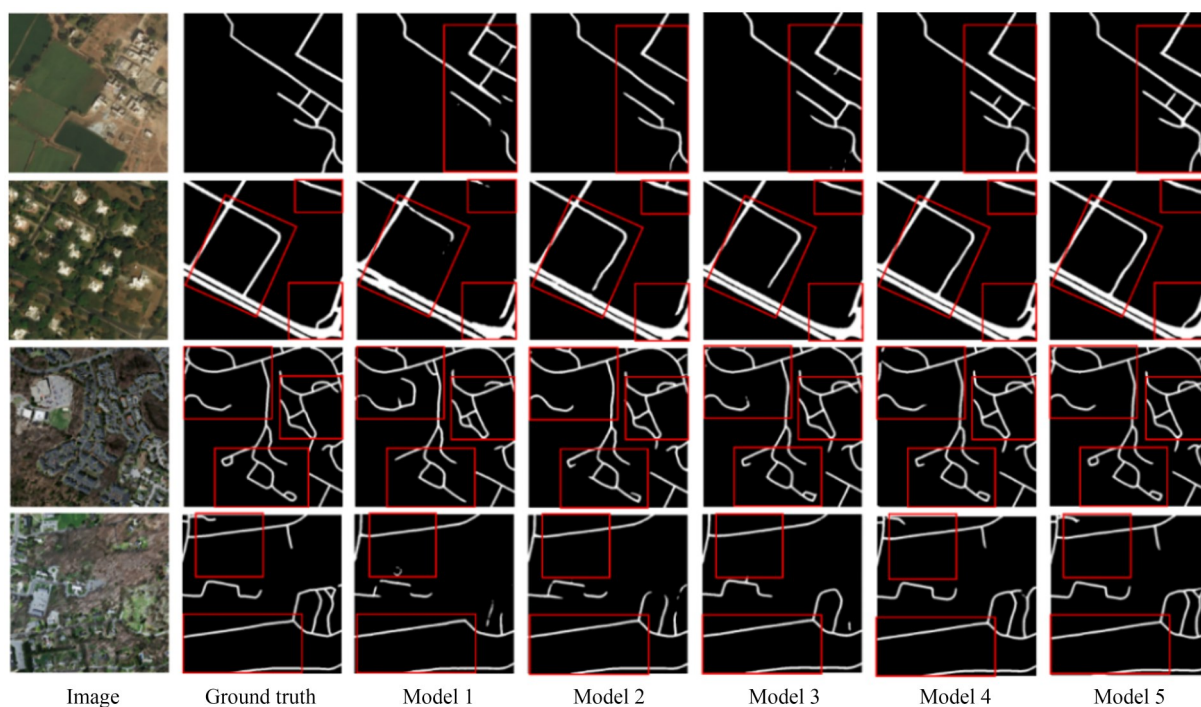


图 9 消融实验可视化结果

Fig. 9 Visualization results of ablation experiment

几何形态,显著提高模型对拓扑错误的敏感度,从而推理出正确的道路拓扑。第二行道路呈现出复杂的几何形态,在添加 L_{TA} 后网络精准恢复出道路拓扑结构。第三行红框中,MDG 模块的引入显著改善断裂现象。第四行红框区域,FGC 引导网络关注模糊的道路特征,有效地抑制环境干扰并缓解了道路细微分支遗漏问题(彩图见期刊电子版)。

DeepGlobe 数据集的定量分析结果如表 7 所

示。Model1 在 Precision, Recall, F1-score 和 IoU 指标分别为 75.08%, 79.11%, 77.04% 与 63.08%;在添加 MDG 模块后,指标分别提高 1.59%, 1.11%, 1.36% 与 2.2%;添加 FGC 模块后,四个指标分别上升 2.10%, 0.71%, 1.44% 与 2.28%;Model4 相较于 Model1 指标分别提高 3.75%, 2.42%, 3.12% 和 3.96%;Model5 与 Model1 相比指标分别提高 5.61%, 4.15%, 4.91%, 6.27%。

表 7 DeepGlobe 消融实验结果

Tab. 7 DeepGlobe ablation experiment results

Model	Baseline	MDG	FGC	L_{BCE}	L_{TA}	Precision/%	Recall/%	F1-score/%	IoU/%
Model1	✓			✓		75.08	79.11	77.04	63.08
Model2	✓	✓		✓		76.67	80.22	78.40	65.28
Model3	✓		✓	✓		77.18	79.82	78.48	65.36
Model4	✓	✓	✓	✓		78.83	81.53	80.16	67.04
Model5	✓	✓	✓		✓	80.69	83.26	81.95	69.35

Massachusetts 数据集的定量分析结果如表 8 所示。Model1 在 Precision, Recall, F1-score 和 IoU 指标分别取得 76.11%, 79.03%, 77.54% 以及 63.18%;在添加 MDG 后指标分别上升 1.12%, 0.39%, 0.77% 和 1.08%;添加 FGC 后指标分别提高 1.51%, 0.09%, 0.82% 和 1.15%;Model4 相较于 Model1, 四个指标分别提

高 2.69%, 0.51%, 1.63% 和 2.61%;Model5 与 Model1 相比, 指标分别提高 3.8%, 1.02%, 2.44%, 3.2%。

综上所述,消融实验结果充分表明,所提模块显著提升分割精度,并在不同数据集上均具有优越的提取效果,表明 RDTA-Net 在光学遥感道路提取任务中具有良好的适应性与鲁棒性。

表 8 Massachusetts 消融实验结果

Tab. 8 Massachusetts ablation experiment results

Model	Baseline	MDG	FGC	L_{BCE}	L_{TA}	Precision/%	Recall/%	F1-score/%	IoU/%
Model1	✓			✓		76.11	79.03	77.54	63.18
Model2	✓	✓		✓		77.23	79.42	78.31	64.26
Model3	✓		✓	✓		77.62	79.12	78.36	64.33
Model4	✓	✓	✓	✓		78.80	79.54	79.17	65.79
Model5	✓	✓	✓		✓	79.91	80.05	79.98	66.38

3.5 损失函数对比实验

为了进一步验证 TALoss 对道路拓扑的约束,本文在 DeepGlobe 数据集上进行详细的对比实验。 L_{BCE} 为二元交叉熵损失, $clDice^{[19]}$ 为拓扑相似性度量, L_{TA} 为本文的 TALoss。

定量分析结果如表 9 所示。在 U-Net 上使用 L_{BCE} , Precision, Recall, F1-score 和 IoU 指标(下文简称为分割指标)分别为 72.70%, 77.66%, 75.10% 与 60.12%, 拓扑分数 β_0 与 β_1 分别取得

1.274 和 1.143 的误差; $clDice$ 与 L_{BCE} 相比分割指标上升 0.86%, 0.83%, 0.85% 与 0.95%, β_0 与 β_1 分别减少了 0.091 与 0.058;采用 L_{TA} 时分割指标提高 1.65%, 1.72%, 1.69% 与 2.20%, β_0 与 β_1 分别降低 0.197 与 0.201。RDTA-Net 采用 L_{BCE} 时, 分割指标分别为 78.83%, 81.53%, 80.16% 与 67.04%, β_0 与 β_1 分别达到 0.708 与 0.691;采用 $clDice$ 时分割指标分别上升 0.86%, 0.83%, 0.85%

与 0.95%, β_0 与 β_1 分别减少 0.065 与 0.074;相较于 L_{BCE} ,采用 L_{TA} 时分割指标分别提高 1.86%, 1.73%,1.79%与 2.32%,并取得最佳分数, β_0 与 β_1 分别降低 0.167 与 0.11,取得最低拓扑误差,代

表预测与真实路网拓扑一致性最高。将 L_{TA} 作为损失时,RDTA-Net 与 U-Net 相比,四个分割指标分别提高 6.34%,3.88%,5.16%,与 7.03%,同时 β_0 与 β_1 的误差值降低 0.536 和 0.361。

表 9 不同损失函数对道路提取的影响
Tab.9 Impact of different loss functions on road extraction

Methods	Loss	Precision/%	Recall/%	F1-score/%	IoU/%	β_0 error	β_1 error
FCN	L_{BCE}	70.53	73.65	72.06	56.32	1.816	1.738
	clDice	71.21	74.30	72.72	56.96	1.715	1.639
	L_{TA}	72.14	75.74	73.90	58.60	1.522	1.396
SegNet	L_{BCE}	69.60	78.17	73.64	58.28	1.584	1.483
	clDice	70.13	78.82	74.22	58.69	1.489	1.317
	L_{TA}	70.92	79.25	74.85	59.26	1.329	1.218
U-Net	L_{BCE}	72.70	77.66	75.10	60.12	1.274	1.143
	clDice	73.56	78.49	75.95	61.07	1.183	1.085
	L_{TA}	74.35	79.38	76.79	62.32	1.077	0.942
PSPNet	L_{BCE}	66.27	79.48	72.28	56.47	1.764	1.681
	clDice	67.06	79.93	72.93	57.43	1.692	1.590
	L_{TA}	68.22	80.52	73.86	58.96	1.451	1.264
RDTA-Net	L_{BCE}	78.83	81.53	80.16	67.04	0.708	0.691
	clDice	79.28	82.37	80.80	67.85	0.643	0.617
	L_{TA}	80.69	83.26	81.95	69.35	0.541	0.581

为验证 TALoss 在实际训练中的可行性,本文在 RDTA-Net 上对比了 BCE, clDice 和 TALoss 的性能与计算效率。为确保实验公平性,所有实验均在相同输入尺寸与参数设置下进行,结果如表 10 所示。实验结果表明,将 BCE 作为损失时,网络运行时间为 0.226 s, IoU 取得 67.04%;clDice 引入形态骨架的相似性度量,提升模型对曲线结构的感知力,使 IoU 提升 0.81%,但由于需额外提取预测与真实骨架,计算时间增加 0.053 s,且骨架提取的不精确影响约束的稳定性;相比之下,TALoss 基于持续同调理论,从拓扑学角度衡量预测与真实路网的拓扑形

态差异,大幅提升分割路网的拓扑正确性,虽然计算耗时增加 0.079 s,但 IoU 提高 1.50%,以略高的计算开销换来了更强的拓扑感知力,展现出优异的实际应用潜力。

综上所述,采用 L_{TA} 可显著增强网络对道路形态的拓扑感知力,使网络更关注于光学遥感道路的曲线特性,并且在像素准确性与拓扑正确性两个方面均具有优越的约束能力,实现了对道路形态的精准推理。

3.6 超参数讨论与分析

图 10 展示了在学习率与批量大小保持一致的条件下不同优化器对模型性能的影响。结果表明,AdamW 优化器表现最佳, IoU 达到 69.35%;其次为 Adam 优化器 IoU 为 69.02%, Nadam 优化器 IoU 为 68.85%,SGD 优化器位列最后, IoU 达到 68.53%。SGD 采用固定学习率,在每次小批量更新后对网络权重进行梯度调整,训练初期收敛速度较慢,且接近局部最优解时易产生波动,导致整体性能提升受限。Nadam 优化

表 10 损失函数计算耗时对比
Tab.10 Time consuming comparison of loss function calculation

Methods	Loss	IoU/%	Training time/s
RDTA-Net	BCE	67.04	0.226
RDTA-Net	clDice	67.85	0.279
RDTA-Net	TALoss	69.35	0.305

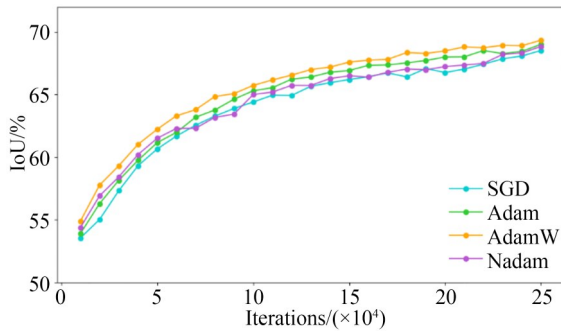


图 10 不同优化器对模型性能的影响

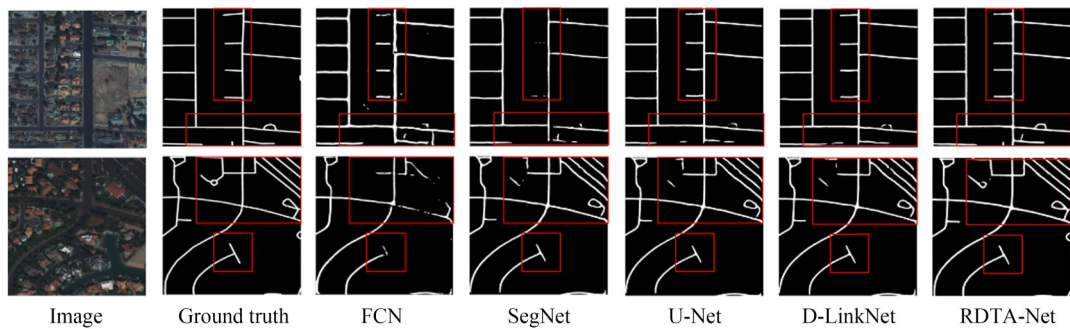
Fig. 10 Influence of different optimizers on model performance

器得益于Nesterov加速,在训练初期收敛速度较快,但因对梯度噪声较敏感,在遥感图像复杂的纹理变化时表现不稳定。Adam优化器的IoU整体稳步上升,即使初期增幅较缓,但凭借自适应学习率机制一定程度提升了训练效率,但因缺乏独立的权重衰减项,导致在正则化方面存在局限性。相比之下,AdamW引入独立的权重衰减机制,使模型训练过程更加稳定。特别在遥感图像分割这种高复杂度任务中,AdamW能够有效缓解过拟合并快速收敛至更优解,因此在本实验中取得了最优性能。

3.7 模型泛化性实验与分析

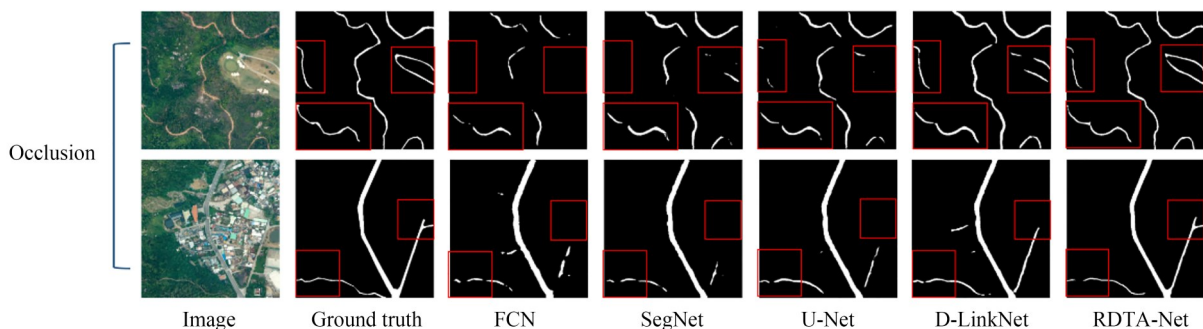
为验证所提网络泛化性,本文在SpaceNet与CHN6-CUG遥感道路数据集上开展了详细的对比实验与鲁棒性分析;并在CHN6-CUG数据集上通过高斯模糊、亮度调整策略进行测试,模拟光学遥感影像中噪声干扰、光照变化等条件,验证所提网络在实际应用中对环境变化的适应能力。测试模型首先在Massachusetts数据集训练到最佳结果,而后进行测试,其中测试影像均未出现在训练集中,且训练区域与测试区域在光照条件、城市规模、空间分辨率等方面存在显著差异。具体来说,训练数据集采集自美国马萨诸塞州,覆盖了城郊区以及农村等多种场景,空间分辨率为1 m/pixel;SpaceNet测试数据包含巴黎、拉斯维加斯、上海和喀土穆四个不同城市,空间分辨率为0.3 m/pixel;CHN6-CUG测试数据涵盖北京、上海、武汉、深圳、香港以及澳门等六个发展程度差异较大的中国城市,空间分辨率为0.5 m/pixel。

各模型的道路提取结果如图11(a)与图11(b)所示。红框区域存在严重的植被遮挡,仅有RDTA-Net准确推理出被遮挡的路段,并保持良



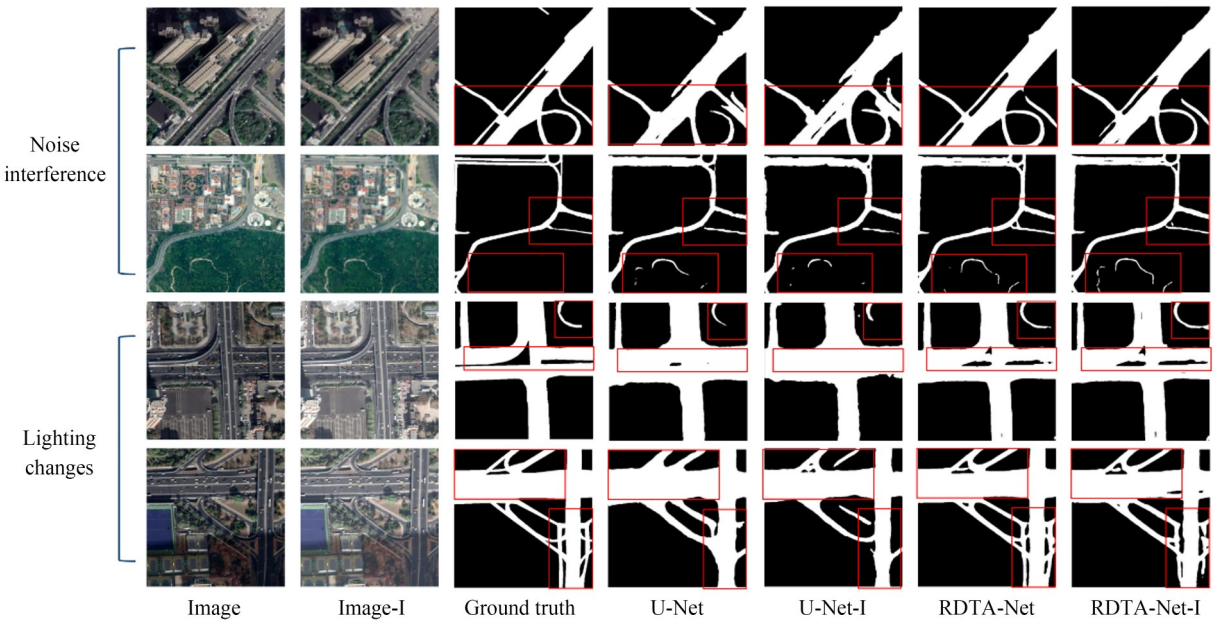
(a) SpaceNet数据集各模型的道路提取结果

(a) Road extraction results of various models in SpaceNet dataset



(b) CHN6-CUG数据集各模型的道路提取结果

(b) Road extraction results of various models in CHN6-CUG dataset



(c) 复杂环境下的道路可视化结果
(c) Road visualization results in complex environments

图 11 泛化性测试结果
Fig. 11 Generalization test results

好的连通性,表明所提模型对道路结构的感知能力较强,在光学遥感影像中具有更好的适应性(彩图见期刊电子版)。

复杂环境下的可视化结果如图 11(c)所示,其中 Image-I 表示添加干扰后的遥感影像,U-Net-I 与 RDTA-Net-I 表示添加干扰后的模型分割结果。前两行为噪声干扰条件,后两行为光照变化条件。在噪声干扰下,所提模型提取的路网更加完整,且与真实路网的几何形态差异较小;

在光照变化条件下,路网连通性优异,几乎没有漏分现象。这表明,RDTA-Net 能有效应对光学遥感影像中的复杂成像条件,提高道路提取的准确性和鲁棒性。

各场景下的定量分析结果如表 11 所示。RDTA-Net 在遮挡、噪声以及光照等复杂环境下,均展现出卓越的鲁棒性和稳定性,验证所提方法在光学遥感道路提取任务中的泛化能力。

表 11 不同场景下的定量实验分析
Tab. 11 Quantitative experimental analysis in different scenarios (%)

Methods	遮挡情况		噪声干扰		光照变化	
	F1-score	IoU	F1-score	IoU	F1-score	IoU
FCN	62.25	48.83	61.14	47.26	61.22	48.07
SegNet	67.19	54.41	66.23	54.03	65.69	53.82
U-Net	68.91	55.02	67.68	54.87	67.34	54.71
D-LinkNet	71.51	57.17	71.25	56.54	70.61	55.58
RDTA-Net	73.42	61.63	72.86	60.38	72.47	60.26

3.8 复杂性分析

为了验证 RDTA-Net 的有效性,本文通过参数量(Params/M)与浮点运算次数(FLOPs/G)

分别评估模型空间、时间复杂度;并通过 FPS、训练时间(Training time/s)两项指标量化模型运行效率。为确保模型复杂度比较的公平性,所有模

型的输入尺寸及参数设置均相同。各模型的复杂度比较结果如表 12 所示。RDTA-Net 与 U-Net 相比,其参数量、FLOPs 与训练时间指标分别增加 16.22 M,43.87 G 与 0.11 s,尽管本文模型的复杂度与时间成本略有增加,但表现出更优异的道路分割性能;相较于 SDUNet,RDTA-Net 在参数量、FLOPs 与训练时间上分别减少 32.98 M,119.82 G 与 0.157 s,表明本方法在实时性方面具有潜在优势,在确保良好连通性的条件下具有更低的复杂度。综上所述,RDTA-Net 在展现出了更优的效率-精度均衡性,以较低的复杂度满足了实际应用中对效率的需求,验证了其在光学遥感道路提取中的卓越性能。

表 12 复杂性分析

Tab. 12 Complexity Analysis

Methods	Params	FLOPs	FPS	Training time
FCN	15.31	80.47	85	0.156
SegNet	18.82	97.11	78	0.173
U-Net	31.04	150.85	64	0.195
PSPNet	53.32	218.61	31	0.379
D-LinkNet	31.1	157.99	69	0.218
SDUNet	80.24	314.54	32	0.462
RDTA-Net	47.26	194.72	53	0.305

4 结 论

针对道路提取结果出现连通性弱、细微分支

遗漏以及与真实路网拓扑结构不一致等问题,本文提出了一种 RDTA-Net 模型用于光学遥感道路分割,该网络通过 MDG 模块关注道路的多方向连通关系,增强道路分支间的联系,提高分割连续性;同时利用 FGC 模块充分挖掘不同粒度特征图的互补特征,强化细节与语义表达,提高网络对细微分支的捕捉能力;最后,为进一步感知道路曲线结构与全局形态,本文利用 TALoss 从拓扑视角探索几何结构差异,增强模型对拓扑变化的敏感度,约束预测路网与真实路网的拓扑一致性。所提模型在 DeepGlobe 与 Massachusetts 数据集上的 F1 值分别为 81.95% 和 79.98%,IoU 分别达到 69.35% 和 66.38%,这两个指标均取得最高分数,这表明 RDTA-Net 在光学遥感道路提取任务中具有优越的提取效果;进一步的泛化性实验结果表明,本方法在遮挡、噪声及光照变化等复杂场景下依然保持明显优势。未来,我们将进一步优化模型的计算效率与拓扑建模精度,以期在更多实际光学遥感应用中实现更高的道路提取性能。

作者贡献声明:

孟月波:分割方法的提出,论文构思和撰写;
黄欣羽:分割方法实验设计、验证与可视化呈现;
苏世龙:可视化呈现,论文审核与编辑写作;
王恒:论文审核与编辑写作。

参考文献:

[1] 戴激光,王杨,杜阳,等. 光学遥感影像道路提取的方法综述[J]. 遥感学报, 2020, 24(7): 804-823.
DAI J G, WANG Y, DU Y, *et al.* Development and prospect of road extraction method for optical remote sensing image[J]. *Journal of Remote Sensing*, 2020, 24(7): 804-823. (in Chinese)
[2] SHAO Y Z, GUO B X, HU X Y, *et al.* Application of a fast linear feature detector to road extraction from remotely sensed imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2010, 4(3): 626-631.
[3] LAPTEV I, MAYER H, LINDEBERG T, *et al.* Automatic extraction of roads from aerial images

based on scale space and snakes[J]. *Machine Vision and Applications*, 2000, 12(1): 23-31.
[4] WEGNER J D, MONTOYA-ZEGARRA J A, SCHINDLER K. A Higher-Order CRF Model for Road Network Extraction[C]. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*. June 23-28, 2013, Portland, OR, USA. IEEE, 2013: 1698-1705.
[5] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation [C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. May 24, 2016, IEEE, 2017: 640-651.
[6] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-

- decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [7] RONNEBERGER O, FISCHER P, BROX T. *U-Net: Convolutional Networks for Biomedical Image Segmentation* [M]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [8] CHAURASIA A, CULURCIELLO E. LinkNet: Exploiting encoder Representations for Efficient Semantic Segmentation[C]. 2017 *IEEE Visual Communications and Image Processing (VCIP)*. December 10-13, 2017, St. Petersburg, FL, USA. IEEE, 2017: 1-4.
- [9] XU Y Y, XIE Z, FENG Y X, *et al.* Road extraction from high-resolution remote sensing imagery using deep learning [J]. *Remote Sensing*, 2018, 10 (9): 1461.
- [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]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. June 18-22, 2018, Salt Lake City, UT, USA. IEEE, 2018: 192-1924.
- [11] YANG Z G, ZHOU D X, YANG Y, *et al.* TransRoadNet: a novel road extraction method for remote sensing images *via* combining high-level semantic feature and context [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 6509505.
- [12] KAMPFFMEYER M, DONG N Q, LIANG X D, *et al.* ConnNet: a long-range relation-aware pixel-connectivity network for salient segmentation [J]. *IEEE Transactions on Image Processing*, 2018.
- [13] MEI J, LI R J, GAO W, *et al.* CoANet: connectivity attention network for road extraction from satellite imagery [J]. *IEEE Transactions on Image Processing*, 2021, 30: 8540-8552.
- [14] JIE Y S, HE H Y, XING K, *et al.* MECA-net: a MultiScale feature encoding and long-range context-aware network for road extraction from remote sensing images [J]. *Remote Sensing*, 2022, 14 (21): 5342.
- [15] 吴强强, 王帅, 王彪, 等. 空间信息感知语义分割模型的高分辨率遥感影像道路提取[J]. 遥感学报, 2022, 26(9): 1872-1885.
- WU Q Q, WANG S, WANG B, *et al.* Road extraction method of high-resolution remote sensing image on the basis of the spatial information perception semantic segmentation model [J]. *National Remote Sensing Bulletin*, 2022, 26 (9) : 1872-1885. (in Chinese)
- [16] HOU Y W, LIU Z Y, ZHANG T, *et al.* C-UNet: complement UNet for remote sensing road extraction[J]. *Sensors*, 2021, 21(6): 2153.
- [17] PANG Y W, LI Y Z, SHEN J B, *et al.* Towards bridging semantic gap to improve semantic segmentation[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2019: 4229-4238.
- [18] MOSINSKA A, MARQUEZ-NEILA P, KOZINSKI M, *et al.* Beyond the pixel-wise loss for topology-aware delineation[C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 3136-3145.
- [19] SHIT S, PAETZOLD J C, SEKUBOYINA A, *et al.* clDice-A novel topology-preserving loss function for tubular structure segmentation [C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25, 2021. Nashville, TN, USA. IEEE, 2021: 16560-16569.
- [20] HU X, LI F, SAMARAS D, *et al.* Topology-preserving deep image segmentation[J]. *Advances in neural information processing systems*, arXiv: 1906.05404.
- [21] WONG C C, VONG C M. Persistent homology based graph convolution network for fine-grained 3d shape segmentation [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, Montreal, QC, Canada. IEEE, 2021: 7078-7087.
- [22] FU J, LIU J, JIANG J, *et al.* Scene segmentation with dual relation-aware attention network[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(6): 2547-2560.
- [23] CARLSSON G. Topology and data [J]. *Bulletin of the American Mathematical Society*, 2009, 46 (2): 255-308.
- [24] EDELSBRUNNER, LETSCHER, ZOMORO-

- DIAN. Topological persistence and simplification [J]. *Discrete & Computational Geometry*, 2002, 28(4): 511-533.
- [25] COHEN-STEINER D, EDELSBRUNNER H, HARER J, *et al.* Lipschitz functions have lp-stable persistence [J]. *Foundations of Computational Mathematics*, 2010, 10(2): 127-139.
- [26] TAHA A A, HANBURY A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool [J]. *BMC Medical Imaging*, 2015, 15: 29.
- [27] ZHAO H S, SHI J P, QI X J, *et al.* Pyramid scene parsing network[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6230-6239.
- [28] DAI L, ZHANG G Y, ZHANG R T. RADANet: road augmented deformable attention network for road extraction from complex high-resolution remote-sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5602213.
- [29] YANG M X, YUAN Y, LIU G C. SDUNet: Road extraction via spatial enhanced and densely connected UNet[J]. *Pattern Recognition*, 2022, 126: 108549.
- [30] LUO L, WANG J X, CHEN S B, *et al.* BDT-Net: road extraction by bi-direction transformer from remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 2505605.
- [31] CHEN J, YANG L B, WANG H, *et al.* Road extraction from high-resolution remote sensing images via local and global context reasoning [J]. *Remote Sensing*, 2023, 15(17): 4177.
- [32] ZHU X H, HUANG X H, CAO W J, *et al.* Road extraction from remote sensing imagery with spatial attention based on swin transformer [J]. *Remote Sensing*, 2024, 16(7): 1183.
- [33] XIONG Y Q, LI L, YUAN D, *et al.* CFRNet: road extraction in remote sensing images based on cascade fusion network [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 6011705.
- [34] LI J, LIU Y, ZHANG Y D, *et al.* Cascaded attention DenseUNet (CADUNet) for road extraction from very-high-resolution images [J]. *ISPRS International Journal of Geo-Information*, 2021, 10(5): 329.
- [35] ZHOU G D, CHEN W T, GUI Q S, *et al.* Split depth-wise separable graph-convolution network for road extraction in complex environments from high-resolution remote-sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5614115.
- [36] WANG Y, PENG Y X, LI W, *et al.* DDU-Net: dual-decoder-U-net for road extraction using high-resolution remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 4412612.
- [37] LI S F, LIAO C, DING Y L, *et al.* Cascaded residual attention enhanced road extraction from remote sensing images [J]. *ISPRS International Journal of Geo-Information*, 2022, 11(1): 9.
- [38] ZHANG R X, ZHU W, LI Y K, *et al.* D-Fusion-Net: road extraction from remote sensing images using dilated convolutional block [J]. *GIScience & Remote Sensing*, 2023, 60(1): 2270806.

作者简介:



孟月波(1979—),女,陕西西安人,教授,西安建筑科技大学信息与控制工程学院硕士生导师,2014年于西安交通大学获得工学博士学位,主要从事计算机视觉理解、建筑环境调控与节能、多模态感知决策方面的研究。E-mail: mengyuebo@163.com