

林雨准,刘智,王淑香,等. 基于卷积神经网络的光学遥感影像道路提取方法研究进展. 吉林大学学报(地球科学版),2024, 54(3):1068-1080. doi:10.13278/j.cnki.jjuese.20230078.

Lin Yuzhun, Liu Zhi, Wang Shuxiang, et al. Research Progress of Road Extraction Method for Optical Remote Sensing Images Based on Convolutional Neural Network. Journal of Jilin University (Earth Science Edition), 2024, 54 (3): 1068 - 1080. doi:10.13278/ j.cnki.jjuese.20230078.

基于卷积神经网络的光学遥感影像 道路提取方法研究进展

林雨准,刘智,王淑香,芮杰,金飞

战略支援部队信息工程大学地理空间信息学院,郑州 450001

摘要:随着光学遥感影像空间分辨率的提升和获取渠道的丰富,利用光学遥感影像实现地物智能解译已成为高效的技术路径。由于卷积神经网络(convolutional neural networks, CNN)强大的特征提取能力以及道路信息在多个领域的应用需求,基于CNN的道路提取方法成为了当前的研究热点。鉴于此,本文根据近年来的相关研究文献,对基于CNN的道路提取方法从基于形状特征的改进、基于连通性的改进、基于多尺度特征的改进和基于提取策略的改进四个方面进行归纳总结,然后描述典型道路遮挡案例,并利用经典CNN从样本标签的局限性层面对当前的技术难点进行分析与验证,最后从多源数据协同、样本库建设、弱监督模型和域适应学习四个方面对遥感影像道路提取的发展趋势进行评估和展望。

关键词:卷积神经网络;光学;遥感影像;道路提取;智能解译

doi:10.13278/j.cnki.jjuese.20230078

中图分类号:P237

文献标志码:A

Research Progress of Road Extraction Method for Optical Remote Sensing Images Based on Convolutional Neural Network

Lin Yuzhun, Liu Zhi, Wang Shuxiang, Rui Jie, Jin Fei

Institute of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China

Abstract: With the improvement of spatial resolution of optical remote sensing images and the enrichment of acquisition channels, optical remote sensing images has become an efficient technological method to achieve intelligent interpretation of land features. Due to the powerful feature extraction ability of convolutional neural networks (CNN) and the demand of road information in many fields, road extraction methods based on CNN have become a current research hotspot. In view of this, this paper

收稿日期:2023-04-02

作者简介:林雨准(1993—),男,博士研究生,讲师,主要从事遥感数据智能解译, E-mail: lyz120218@163.com

通信作者:金飞(1984—),男,副教授,博士,主要从事遥感数据智能处理与分析, E-mail: jf371@sina.com

基金项目:国家自然科学基金项目(42201443)

Supported by the National Natural Science Foundation of China (42201443)

summarizes the road extraction method based on CNN from four aspects: Improvement of shape features, improvement of connectivity, improvement of multi-scale features and improvement of extraction strategy according to the relevant research literature in recent years. Then, we describe typical road occlusion cases and use classical CNNs to analyze and validate the current technical difficulties at the level of limitations of sample labels. Finally, the development trends of road extraction from remote sensing images are outlooked from four aspects, namely, multi-source data synergy, sample library construction, weakly supervised modeling and domain-adaptive learning.

Key words: convolutional neural network; optical; remote sensing image; road extraction; intelligent interpretation

0 引言

智能化、自动化、实时化的道路提取技术在地图更新、自动驾驶、城市规划、车辆导航和应急响应等领域均具有重要作用。光学遥感影像(后简称为“影像”)作为非接触信息获取技术的产品,其传播过程的电磁波信息可与地物类别形成规律的映射关系,且当前影像的空间分辨率、光谱分辨率和时间分辨率不断提升,使其在宏观性、直观性、实时性等方面的优势进一步突显。此外,道路作为核心的交通设施,由于其现实空间位置的表层化,使其往往可以在影像中展现光谱信息、纹理规律、形状特征和拓扑结构。因此,影像可以为大范围的道路提取提供有效的数据支持。

关于影像的道路提取方法,根据其发展过程主要可以分为传统方法和深度学习方法。传统方法主要包括模板匹配法^[1-3]、知识驱动法^[4-5]和面向对象法^[6-7]等,该类方法主要是依托道路在影像上呈现的形状特征、光谱特征、纹理特征以及人为设计的浅层组合特征展开的;但是随着影像空间分辨率的不断提升,道路在浅层特征空间中往往不够聚集并且与其他地物存在交集,导致该类方法构建的道路提取模型无法匹配道路特征的复杂性和不确定性。关于该类方法的研究现状在文献^[8-11]中均有论述(本文不再详细介绍),本文将着重对深度学习方法展开讨论。

近年来,卷积神经网络(convolutional neural networks, CNN)不断发展,在目标检测^[12]、数据去噪^[13]等领域的应用愈发广泛,利用卷积、池化、归一化、非线性激活等操作提取的深层特征相比浅层特征在相应特征空间内,区分地物类别的依据更加明确和清晰,有效提升了计算机影像智能处理与解译

的稳定度、准确性和效率。尤其诸如 FCN (fully convolutional networks)^[14]、U-Net^[15]、SegNet^[16]和 DeepLab 系列^[17-20]等经典语义分割网络的诞生,使得端到端 CNN 在像素级智能解译中得到了广泛应用,相关技术在道路提取(二值语义分割)中也不断地探索进步。比如早期文献^[21-23]通过在特征提取阶段引入残差模块^[24]、在损失函数计算阶段引入正负样本平衡机制等方式提升道路提取结果的稳定性和精度,但对道路本身的属性特征挖掘有限。目前,考虑到道路提取面向的困难以及道路的形状和尺度规律等,学者们主要在典型语义分割网络框架的基础上围绕道路的形状、连通性、尺度和整体的提取策略实施改进,故本文将对上述改进进行总结归纳,并对当前的技术难点和发展趋势进行评估和展望。

1 研究现状

基于 CNN 的道路提取基本流程如图 1 所示,主要包括训练和测试两个阶段。在训练阶段,首先进行前向传播,即将影像(训练集)送入 CNN 获得预测结果;然后通过损失函数计算预测结果和道路标签的相似程度或者差异程度,接着本着差异最小原则,通过后向传播对 CNN 模型的参数进行更新;最后不断地重复前三个步骤,直到重复次数达到设定阈值,或预测结果和标签的差异不再减小。这个阶段总体来说就是通过训练集的影像和对应的道路标签对 CNN 模型的参数进行更新,使得预测结果尽可能接近真值。第二个阶段是测试,即将待提取道路的影像输入到已训练好(模型参数已更新)的 CNN 模型中,通过生产预测(前向传播)的路线完成道路提取。因此,基于 CNN 的道路提取技术涉及的核心要素就是 CNN 模型,该模型的设计也是影

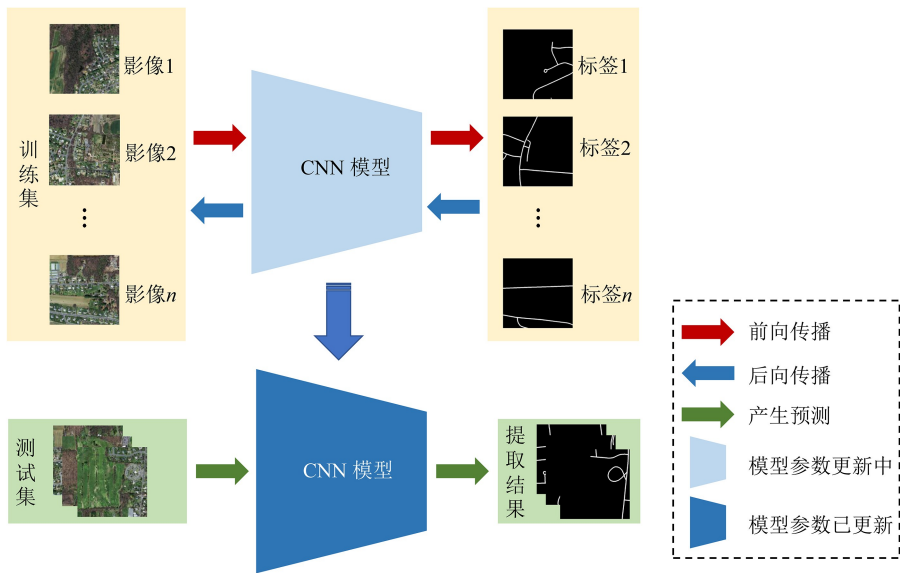


图 1 基于卷积神经网络的道路提取基本流程

Fig.1 Basic process of road extraction based on convolutional neural network

像语义分割的关键。考虑到道路形状、提取结果面向的应用需求等特殊特性,针对影像道路提取问题,当前学者在典型语义分割网络的框架下不断进行了改进,主要包括基于形状特征的改进、基于连通性的改进、基于多尺度特征的改进以及基于提取策略的改进。

1.1 基于形状特征的改进

不同于建筑物、植被或湖泊等地形要素,道路在影像上往往呈现条状,这样的规律在道路提取中可以作为辅助条件,提升道路提取结果的可靠性。目前,依据道路形状的改进主要包括卷积核和标签数据两个方面。关于卷积核,主要是设计与道路形状相适应的卷积窗口,从而对特征提取的方向和空间进行约束,如文献[25-27]在 CNN 模型的上采样中均采用 0° 、 45° 、 90° 和 135° 四个方向的条状卷积进行特征提取。从形状匹配度来看,上述文献设计的条状卷积核比常规的方形卷积核更适合道路提取。方形卷积核对周围各方向的像素均有涉及,虽然卷积核的参数可以根据相关程度进行调整,但不可避免地会较多地吸收相邻像素的无关信息;而条状卷积核沿着空间方向的长条形状来捕获道路的相关信息,在形状特征层面防止了不相关像素的干扰。

关于标签数据方面,主要是根据已有的道路标签为目标赋予相应的方向属性,生成方向标签。如Batra等[28]基于道路的形状提出了“方向学习”的概念,具体操作是给图像中的每个道路像素赋予真实走向的矢量信息,通过真值矢量和道路标签共同监督道

路提取模型的学习。Zhou等[29]将道路的方向平均分成36个等份,每一部分为 10° ,因此每个像素在方向学习的这条路线上输出36维的有效结果,即同一个位置的不同通道表示属于对应方向的概率。上述两篇文章的方向学习都是基于矢量形式获取监督信息的,文献[30]直接基于栅格数据展开方向学习,以像素级的道路提取结果为基准,首先判断像素是否属于道路,若不属于道路则不赋予方向属性,若属于道路则在一定圆形缓冲区内确定每个方向上属于道路的像素数目,最后将数目最多的方向作为该像素的方向标签,其核心思想为获取二值图像的主方向。

形状特征模块充分考虑了道路在影像上呈现的态势规律,将常规的特征提取或者标签数据在重点方向、重点区域进行聚焦,相关文献[25-30]证明了该改进对条状类目标的适用性,其提取结果在像素级和连通性方面均有一定提升。

1.2 基于连通性的改进

道路作为重要的交通运输设施,其连通性的准确率对实施方法能否满足应用需求是至关重要的。对于道路而言,正确的连通性代表着正确的路线,一旦路线错误会导致更长的行驶时间,或者进入限制区域,甚至带来交通事故等一系列问题。目前关于道路连通性的改进主要包括标签数据和损失函数两个方面。关于标签数据,主要是以像素级标签为基础,构建体现道路拓扑结构的连通性标签。如Mei等[26]基于像素级的道路真值,对当前点八个方向的

邻域依次进行判断。判断的内容为当前点和邻域点是否均为道路,若符合条件,则将当前点在对应方向的连通性标签上作为前景,否则作为背景。以此类推,逐像素逐方向地遍历所有待判断点位,最终生成八通道的连通性标签。另外,Li等^[31]也采取了类似的策略生成连通性标签,不同之处在于标签的数值表示的是八邻域空间内属于道路的像素总数目,故经过此过程后将分类问题转换为回归问题。

除了连通性标签外,部分学者从损失函数的角度对提升道路的连通性展开了研究,关于损失函数对影像道路提取的影响,在文献^[32]中有相关论述。考虑到道路提取属于二分类问题,常用的交叉熵损失函数等虽然可以反映提取结果与真实结果的差异,但计算差异的过程大多是将道路视作点目标或者一定范围的面目标,没有考虑线状地物本身的连通关系,故上述方式计算的损失值很难体现道路拓扑结构的恢复程度,造成道路提取结果与真实的应用场景需求差距较大。基于上述分析,Mosinska等^[33]利用在ImageNet数据集上预训练好的VGG19模型^[34]对预测结果和道路标签分别进行深层特征提取,然后基于深层特征进行损失函数计算,通过实验证明了这样的损失函数突出了线性结构在深层视觉特征中的差异。另外,Oner等^[35]利用背景区域之间产生的分裂来表达道路的连通性,即设计损失函数防止背景区域之间的不必要连接,从而缩小预测道路与真实道路在拓扑结构上的差异。

连通性作为道路的重要属性,其规律在影像中主要体现在邻域信息内,相关文献^[26, 31-33, 35]分别从标签数据和损失函数两个层面进行了邻域范围的关联,实际上是高分辨率影像处理单元从像素到对象的体现,符合影像本身的发展规律。

1.3 基于多尺度特征的改进

在地形图中,道路(公路)包括高速公路、普通公路和简易公路等,不同等级的道路在影像上体现的大小特征往往存在差异。因此,对影像进行道路提取时,需要同时面对尺度不一的对象。CNN统一从小范围的局部特征逐渐扩展到大范围的局部特征,但是在特征逐渐深入、感受野不断扩大的过程中,对关注目标始终是按照统一尺度的处理单元具体实施的,无法兼顾目标的尺度差异。为使模型能同时捕获多尺度特征,学者们展开了相关研究。如Guo等^[36]在编码层中提出了多尺度卷积注意力,在经过 5×5 卷积后分成了四个分支,第一个分支不作任何

处理,第二个至第四个分支分别由长度为7、11、21的水平条状卷积核和垂直条状卷积核进行不同尺度的特征提取,然后将四个分支的结果相加实现多尺度特征融合。Gao等^[37]则对常规的残差卷积模块进行改进,即在 1×1 卷积核之后进行通道分离,然后依次进行 3×3 卷积,由于叠加的 3×3 卷积将扩大感受野,即可借此实现不同尺度的特征提取。

除了特征提取的尺度差异外,考虑到CNN在全局特征捕获上的劣势,部分研究从兼顾全局特征和局部特征的角度实现了多尺度的特征融合。如Zhou等^[38]在编码层和解码层之间引入了图神经网络,有效地结合了CNN在局部特征提取以及图神经网络在全局特征提取中的优势;又如文献^[39-40]在编码层和解码层之间引入了多尺度空洞卷积,实现了不同尺度的特征提取。此外,文献^[41]为进一步突出空间和通道的有效区域,将多尺度空洞卷积和空间通道注意力机制^[42]相结合,文献^[26, 43]则利用了空洞空间卷积池化金字塔^[19],其目的均在于网络模型能够兼顾局部和全局两种层次的特征。

大尺度道路是交通网络的骨干核心,小尺度道路是交通网络保持完整性的必备要素。因此,无差别的提取能力是道路提取方法是否有效的重要评价指标。相关文献^[26, 36-43]从卷积核的尺寸以及利用图神经网络、多尺度空洞卷积等角度进行了探索,取得了一定成效。

1.4 基于提取策略的改进

在语义分割中,网络模型一般采取“编码—解码”的基本框架,但是单一的流水线式网络框架的稳定性和适用性往往有限。因此,学者们从提取策略的角度对方法进行了改进,主要包括三个方面:一是多任务的提取策略,二是由粗到精的提取策略,三是多方法的提取策略。

多任务的提取策略主要是在像素标签的基础上引入其他标签数据对网络共同进行监督,如文献^[31, 44-46]都是采取了这种策略。以文献^[44]为例,所提方法一共包括三个子网络和两个阶段:第一个阶段构建的子网络利用像素级标签进行监督;第二个阶段是在第一个阶段完成后进行的,两个分支的子网络结构相同,输入数据均是原始影像和第一个阶段预测结果的叠加(与由粗到精的提取策略相似),监督数据分别是道路边缘标签和道路中心线标签,从而充分利用道路边缘和中心线在轮廓范围和拓扑结构上的约束优化最终的道路提取结果。

由粗到精的提取策略是采用两个网络先后对影像进行道路提取,其中粗提取的输出一般会作为精提取的输入,从而有效叠加不同网络的提取效果。如 Ding 等^[30]在粗提取阶段将原始影像输入网络,经过结构标签、方向标签和像素级标签共同监督生成粗提取结果;在精提取阶段,将粗提取结果产生的概率图输入至相应网络,直接由像素级标签获得精提取结果。Zhou 等^[29]的策略与 Ding 等^[30]基本一致,除去网络结构本身的差异外,精提取的输入不仅包含粗提取结果,同时还包括粗提取在编码过程中产生的浅层特征。

多方法的提取策略主要是将 CNN 与其他方法相结合,比如 Zhou 等^[38]联合 CNN 与图神经网络就是应用案例之一。另外,文献[40, 47-48]也都采取了多方法的提取策略,其中:文献[47]的整体思想是将强化学习与 CNN 相结合,实施的形式是强化学习策略,即利用节点跟踪进行道路提取,而决策的过程是借助 CNN 完成的;文献[40, 48]则将文献[47]的方法和 CNN 进行结果级融合,进一步提升了提取结果的可靠性。

多任务、多方法或者由粗到精等提取策略的目的在于弥补单一手段在道路提取中可能存在的不足,与多源数据类似,是结合各方面的优势,提升道路提取结果在各类情形下的鲁棒性。这样的技术方法也是解决道路所处环境存在的目标复杂度大、区域差异性大等困难的手段之一。

1.5 方法小节

基于 CNN 的道路提取方法凭借卷积、池化、非线性激活等基本操作实现道路与非道路的特征空间分离。基于 CNN 的道路提取方法和传统方法的总体性能比较情况如表 1 所示。可以看到,传统方法的模型复杂度较低、预处理时间较短、对硬件的要求不高,但是对复杂背景的适应能力和应用时效性较

差,处理过程需要过多人为干预,故该方法在智能解译的自动化层面相对滞后。尤其随着当前影像空间分辨率来到亚米级时代,其对地物的细节还原能力日益突出,传统方法的发展受到了较大影响。但是,传统方法的逻辑基础和理论知识与基于 CNN 的道路提取方法存在一定的共通性,故对传统方法的学习将有助于理解道路在影像上的特征表现规律和不同层次特征之间的关系。

CNN 的四种改进是学者们根据道路在影像中的表现规律以及提取结果的应用需求展开的。若从 CNN 的具体流程出发,四种改进类型又可以分为网络组成单元的改进、网络结构框架的改进和损失函数的改进,其中网络组成单元与结构框架的改进会充分考虑道路的自身属性,但也会增加网络复杂度,降低整体效率;而损失函数的改进摆脱了训练的推理过程,故可以在不同的网络模型间直接套用。总体来说,为最大程度地适应道路的形状、尺度和连通性特征及其在高分辨率影像上的表现规律,目前的方法大多会同时进行多个方面的改进,以期实现精度、效率和稳定性的全面提升。

2 技术难点

目前,随着计算机硬件和网络模型的不断进步,基于 CNN 的道路提取方法在时效性和智能化程度上有了显著的进步,但距离实际应用仍具有一定差距,具体主要体现在影像数据和样本标签两个方面。

2.1 影像数据的局限性

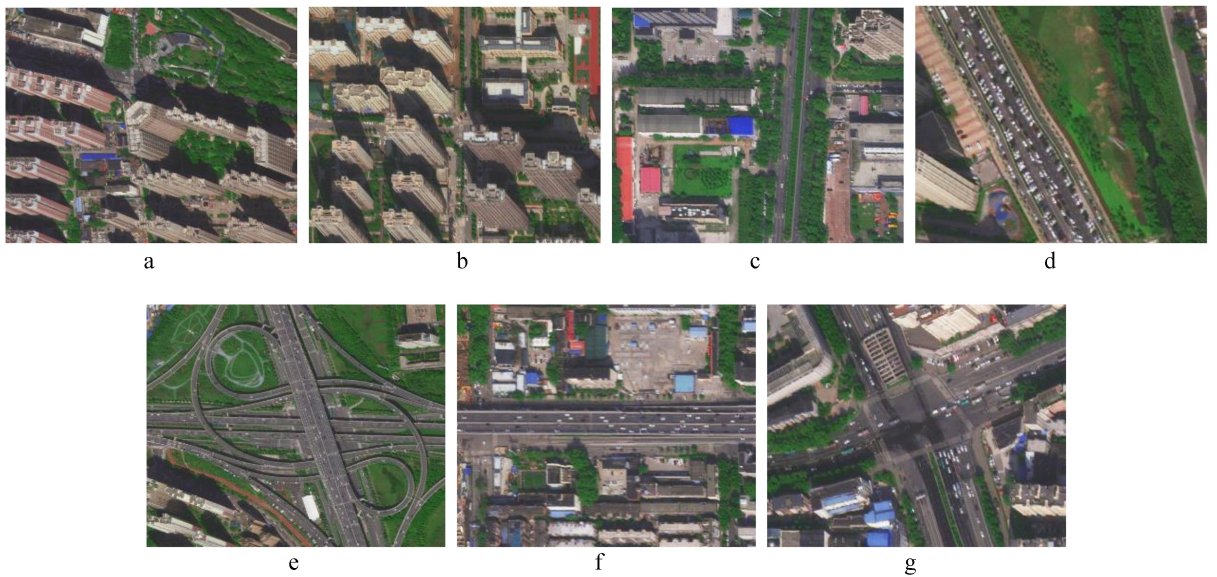
影像数据的局限性主要是由可见光范围内的电磁波穿透能力不足,以及地物分布规律的复杂性等因素造成的。如图 2 所示,在影像中的道路,会因为高大目标的遮挡造成道路反射的电磁波信息无法被遥感器接收从而形成阴影,卫星姿态和光线入射角度造成高大建筑物的顶部信息出现在道路表面从而

表 1 传统方法和卷积神经网络总体性能比较统计

Table 1 Overall performance comparison statistics between traditional method and convolutional neural network method

方法	模型复杂度	模型迁移能力	硬件要求	适用范围	预处理时间	应用时效性
传统方法	较低*	不同数据的关键参数往往需要人工干预	较低*	背景相对简单	较短*	与算法复杂程度相关
卷积神经网络	较高	模型参数可自行学习*	较高	背景相对复杂*	模型训练时间较长	模型训练完成后的推理时间较短*

注:加*表示优势项。



a. 阴影压盖;b. 建筑物遮挡;c. 植被遮挡;d. 活动目标压盖;e. 立交桥连接处压盖;f. 上下行道路压盖;g. 地下隧道。

图 2 道路压盖现象示意图

Fig.2 Schematic diagram of road capping phenomenon

形成遮挡,以及植被、车辆甚至是不同方向、不同等级的道路在空间上重合而形成压盖。上述原因均会导致影像中应属于道路的像素表现形式变为其他地物,从而造成道路提取结果的不完整、不连续和拓扑结构的不准确等问题。

此外,地物分布规律的复杂性体现在道路(公路)与铁路、沟渠、河流的相似性,道路尺度的多样性,道路在高分辨率影像中的车道线、交通标示线,以及道路表面的材质差异等引起的噪声等。以上问题在基于 CNN 的方法中往往会造成模型的欠拟合,或者对某地区的道路过拟合等问题。

2.2 样本标签的局限性

目前,随着深度学习技术的不断发展和智能化

应用需求的不断推进,基于深度学习的影像道路数据集也越来越丰富。表 2 为当前几种典型影像道路提取数据集的基本情况,可以看到目前数据集涵盖的区域和场景较为多样,并且空间分辨率基本都达到了米级、亚米级,这为实现多地区多类型的道路自动智能提取提供了很好的训练样本。

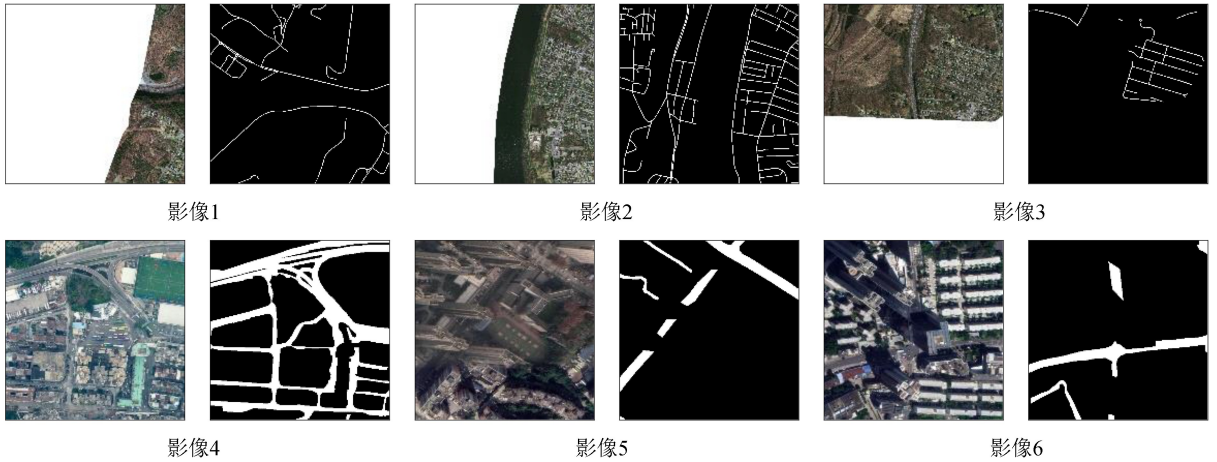
然而,部分数据集提供的参考标准本身存在一定的干扰。如图 3 所示,以 Massachusetts 数据集和 CHN6-CUG 数据集为例:Massachusetts 数据集发布者对训练集中的部分影像随机删除了部分区域的光谱信息,使之成为了干扰项,这也考验了网络模型的泛化能力;CHN6-CUG 数据集的部分标签数据存在锯齿和间接性中断现象,但通过肉眼观察可

表 2 典型影像道路提取数据集

Table 2 Typical image road extraction dataset

数据集名称	训练集数量/幅	验证集数量/幅	测试集数量/幅	区域/场景类型	空间分辨率/m
DeepGlobe ^[49]	6 226	1 243, 无标签	1 101, 无标签	泰国、印度、印度尼西亚/城区、乡村、荒郊	0.5
Massachusetts ^[50]	1 108	14	49	美国马萨诸塞州/城区	1.2
SpaceNet ^[51]	2 780	—	—	拉斯维加斯、上海、巴黎、喀土穆/城区、乡村、荒郊	0.3
CHN6-CUG ^[52]	3 608	—	903	北京、上海、武汉、深圳、香港、澳门/城区	0.5
LRSNY ^[53]	716	220	432	美国纽约	0.5
RoadNet-Data ^[44]	8 km ²	—	—	加拿大渥太华/城区、乡村	0.21

注:“—”表示训练集、验证集和测试集在发布时未明确划分。



影像 1、影像 2、影像 3. Massachusetts 训练集;影像 4、影像 5、影像 6. CHN6-CUG 训练集。

图 3 道路提取训练集干扰示例

Fig.3 Example of interference in the training set for road extraction

以发现,上述现象覆盖的区域由于投影误差等原因造成了影像同一位置的地物叠加问题,这也是影像数据局限性有待解决的技术难点。

此外,目前道路样本的丰富度还无法匹配道路在全球范围内的复杂度,且常规的深度学习模型在样本集与测试集之间的跨域能力还不够理想。以 U-Net^[15]、D-LinkNet^[39] 和 DeepLabv3+^[20] 在 DeepGlobe 和 Massachusetts 两个数据集的提取实验为例,DeepGlobe 数据集涵盖区域的场景类型相对丰富,空间分辨率也更高(表 2)。表 3 是两种数据集交叉训练测试的精度统计结果,五个评价指标

正确率(P)、召回率(R)、 F_1 分数、总体精度(A_o)和交并比(r_{IoU})的计算公式如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}; R = \frac{N_{TP}}{N_{TP} + N_{FN}};$$

$$F_1 = 2 \frac{P \cdot R}{P + R}; A_o = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FP} + N_{TN} + N_{FN}};$$

$$r_{IoU} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}。$$

式中, N_{TP} 、 N_{FP} 、 N_{TN} 、 N_{FN} 分别为正确提取正样本、错误提取正样本、正确提取负样本和错误提取负样本的数量。通过表 3 可以看到,当训练集和测试集

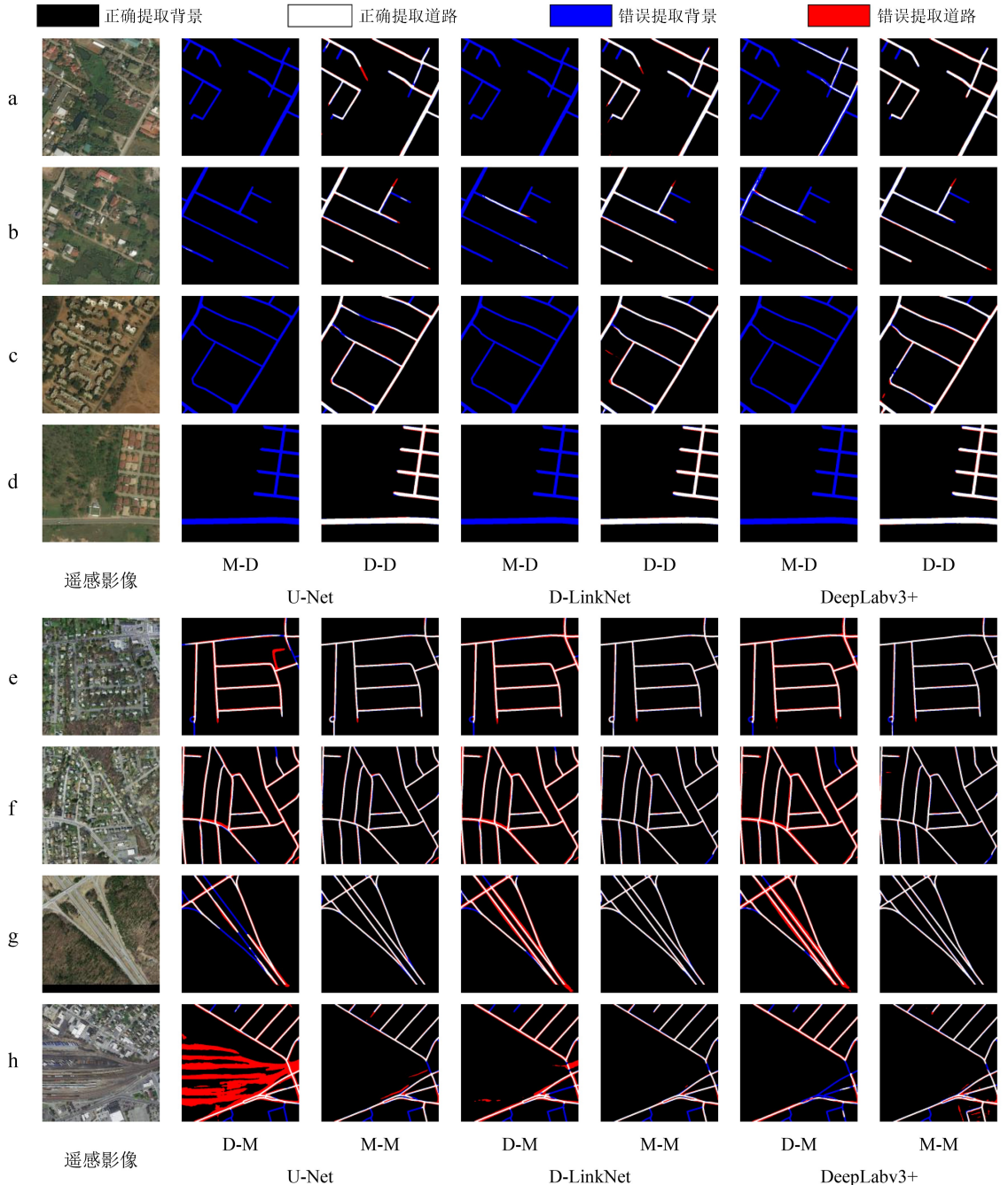
表 3 不同数据集测试结果精度统计

Table 3 Accuracy statistics of test results for different data sets

训练集	测试集	方法	$P/\%$	$R/\%$	$F_1/\%$	$A_o/\%$	$r_{IoU}/\%$
Massachusetts	DeepGlobe	U-Net	68.38	2.19	4.24	95.79	2.17
		D-LinkNet	68.11	3.09	5.91	95.81	3.04
		DeepLabv3+	73.18	3.59	6.84	95.84	3.54
DeepGlobe	DeepGlobe	U-Net	79.88	76.98	78.40	98.19	64.48
		D-LinkNet	82.06	78.39	80.18	98.35	66.92
		DeepLabv3+	82.24	80.09	81.15	98.42	68.28
DeepGlobe	Massachusetts	U-Net	57.06	55.52	56.28	95.73	39.16
		D-LinkNet	58.80	64.55	61.54	96.01	44.44
		DeepLabv3+	59.76	62.72	61.20	96.07	44.10
Massachusetts	Massachusetts	U-Net	78.73	75.79	77.23	97.78	62.91
		D-LinkNet	80.71	76.37	78.48	97.93	64.58
		DeepLabv3+	81.17	74.98	77.95	97.90	63.87

是同一个数据集时各项精度指标明显更高;说明即使 CNN 可以进行深层特征的挖掘,但道路在不同的区域和场景下,仍无法在相应的特征空间有效聚集。因此,要想借助 CNN 实现某一地区的高质量道路提取,样本的选择对提取效果会有明显影响。其次,当训练集是 Massachusetts 数据集、测试集是

DeepGlobe 数据集时的精度指标 R 几乎为 0,结合图 4 可以发现,此时道路提取能力很低。同时,通过图 4 可以看到,D - M 的道路提取效果明显优于 M - D,通过分析两种数据集不难发现,Massachusetts 数据集范围集中在同一区域,影像的地物分布,尤其是道路影像特征相对单一,故其构



M - D 表示训练集是 Massachusetts,测试集是 DeepGlobe;M - M,D - M,D - D 类似。

图 4 道路跨域提取结果示例

Fig.4 Example of road cross-domain extraction results

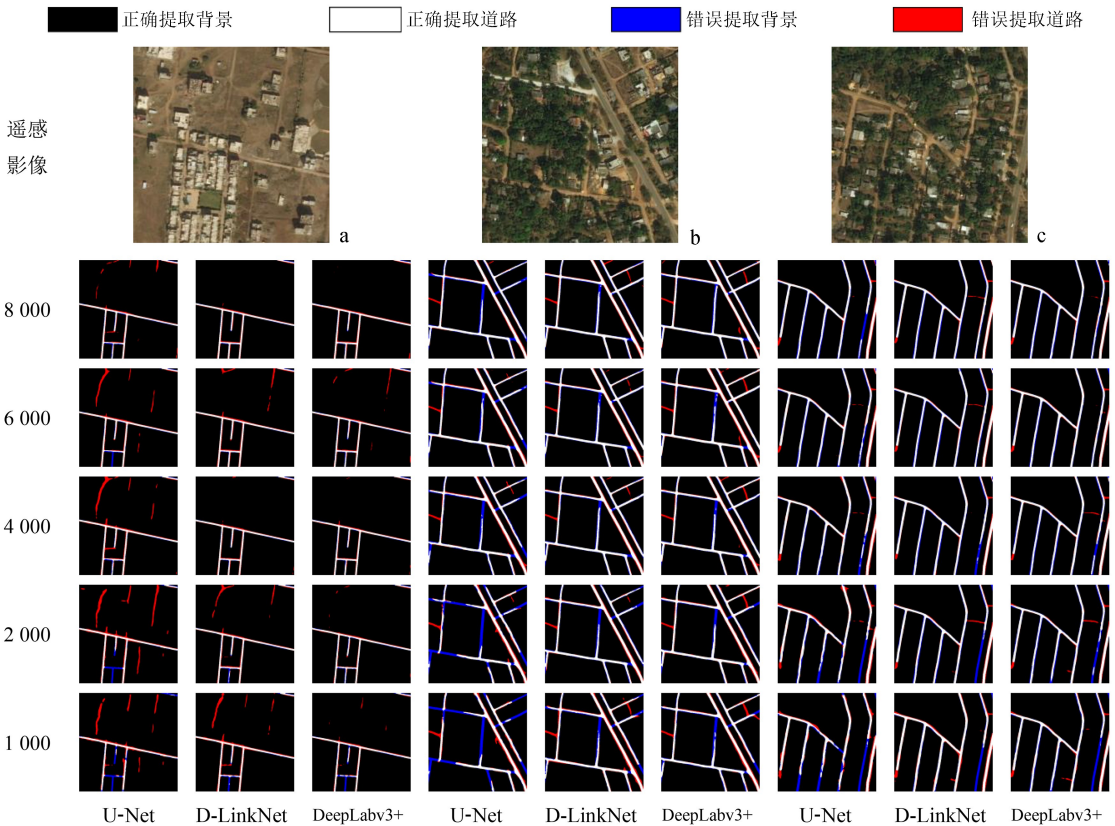
建的模型参数更容易出现过拟合问题,导致网络模型的适应能力较弱。综上所述,样本的质量对构建模型的适用性是至关重要的,但像素级的道路标签需要消耗大量的时间和人力,其成本与智能化的目的相违背。

除样本数据的跨越问题外,样本的数量同样对常规网络模型的有效性有重要影响。以 U-Net、D-LinkNet 和 DeepLabv3+ 在 DeepGlobe 数据集上的提取实验为例,分别在样本数量为 8 000、6 000、4 000、2 000、1 000 的条件下进行道路提取,其中为保证实验结果不受训练影像本身特征的影响,确定训练集时依次从相邻更大数量的训练集中随机抽取(如数量为 4 000 的训练集从数量为 6 000 的训练集中抽取,数量为 8 000 的训练集从原始数据集中抽取)。五种样本数量下的道路提取结果精度统计如表 4 所示,提取结果示例如图 5 所示。结合表 4 和图 5 可以看到,随着样本数量的降低,各项精度指标

表 4 不同样本数量道路提取结果精度统计

Table 4 Accuracy statistics of road extraction results with different sample sizes

方法	样本数量	P/%	R/%	F ₁ /%	A _o /%	r _{10U} /%
U-Net	8 000	78.62	71.85	75.08	97.97	60.10
	6 000	76.35	71.62	73.91	97.85	58.62
	4 000	75.95	69.64	72.66	97.77	57.06
	2 000	69.79	61.23	65.23	97.22	48.40
	1 000	67.95	57.15	62.08	97.03	45.02
D-LinkNet	8 000	80.13	77.05	78.56	98.21	64.69
	6 000	79.27	75.81	77.50	98.13	63.27
	4 000	78.43	74.44	76.38	98.04	61.79
	2 000	77.09	70.56	73.68	97.85	58.33
	1 000	72.88	65.62	69.06	97.50	52.74
Deep-Labv3+	8 000	79.74	77.99	78.86	98.22	65.09
	6 000	79.70	76.05	77.83	98.16	63.71
	4 000	80.02	73.62	76.69	98.09	62.19
	2 000	77.01	70.79	73.77	97.86	58.44
	1 000	75.68	66.10	70.57	97.65	54.52



8 000、6 000、4 000、2 000、1 000 表示训练集的样本数量。

图 5 不同样本数量的道路提取结果示例

Fig.5 Example of road extraction results for different sample sizes

逐渐下降,并且下降的速度在样本数量较小时更加明显,最终的提取结果在道路漏提取上更加显著,尤其当样本数量为1 000、2 000时,道路的提取结果与真实情况差异很大。然而,样本如果是通过人工标注获得,其数量越大往往意味着样本库建立涉及的人工干预越多。因此,如何在小样本条件下保持道路提取结果的精度也是当前的技术难点。

3 展望

影像道路提取技术的发展趋势基本是围绕着影像数据的局限性和样本标签的局限性两个技术难点展开的。关于具体的发展趋势本文主要从以下四点进行展望。

1)多源数据协同的道路提取技术:影像的单视角在信息维度上是存在缺陷的,故引入多源数据进行特征维度的扩充是目前的发展趋势,如Li等^[54]将影像和轨迹数据的特征图叠加共同作为CNN模型的输入,从而在特征层次将影像的静态光谱信息和轨迹数据的动态先验知识有效结合,进一步增加了道路与非道路之间的区别度。另外,街景图像^[55]、点云数据^[56]等不同层次的协同作业将有助于道路在不同维度的特征提取,推动道路提取方法由理论研究走向实践应用。

2)历史数据驱动的道路样本库:依托人工作业进行逐像素或逐区域的样本标记显然与道路提取的智能化发展相矛盾,故借助历史数据有效解决监督类地物提取中面临的样本压力是全面实现道路提取自动化的关键步骤。事实上,道路作为地形要素的核心,在社会的生产作业中往往会留下痕迹,比如浮动车轨迹数据^[57-58]、历史地形图矢量数据等,这些数据的处理分析将有助于实现道路样本的自动获取。

3)稀疏样本条件下的弱监督模型:由于数据本身的质量差异及其与影像的匹配精度,以及样本在监督学习中的数量和质量要求,自动获取的样本往往无法直接满足其应用需求。因此,从技术手段方面来说,目前主流的发展趋势是弱监督学习^[59-60]的思想,如利用线注释、点注释等稀疏样本标签直接进行网络训练。另外,生成对抗网络也是解决样本质量问题的技术手段。

4)域适应支持下的迁移学习能力:CNN作为监督学习的一种形式,是一种数据驱动技术,其泛化性能受到训练样本的约束。常规模型的迁移学习能力

有限,而不同地区的道路结构、建设布局以及获取影像的光照条件和传感器性能差异都会对常规模型的推广应用带来挑战。域适应作为不同数据集之间的连接工具,有助于提升网络模型的泛化能力,也是道路提取技术摆脱样本库约束的有效手段。因此,提升模型的迁移学习能力有助于道路提取技术的普及和应用范围的扩大。

4 结语

本文以卷积神经网络为技术核心,针对光学遥感影像道路提取方法的研究现状展开了系统梳理与分析,并围绕影像数据局限性和样本标签局限性进行了经典卷积神经网络的定量精度评价与定性效果分析,从而引出未来的发展趋势。总体而言,与其他遥感影像智能解译任务相似,针对道路本身的特征挖掘手段已相对丰富,但如何充分利用现有的数据优势,建立多源数据统筹的道路提取范式是接下来的研究重点和技术趋势。

参考文献(References):

- [1] Leninisha S, Vani K. Water Flow Based Geometric Active Deformable Model for Road Network [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 102: 140 - 147.
- [2] Fu G, Zhao H R, Li C, et al. Road Detection from Optical Remote Sensing Imagery Using Circular Projection Matching and Tracking Strategy[J]. Journal of the Indian Society of Remote Sensing, 2013, 41(4): 819 - 831.
- [3] 林祥国,张继贤,李海涛,等.基于T型模板匹配半自动提取高分辨率遥感影像带状道路[J].武汉大学学报(信息科学版),2009,34(3):293-296.
Lin Xiangguo, Zhang Jixian, Li Haitao, et al. Semi-Automatic Extraction of Ribbon Road from High Resolution Remotely Sensed Imagery by a T-Shaped Template Matching [J]. Geomatics and Information Science of Wuhan University, 2009, 34(3): 293 - 296.
- [4] 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.
- [5] Treash K, Amaratunga K. Automatic Road Detection in Grayscale Aerial Images[J]. Journal of Computing in Civil Engineering, 2000, 14(1): 60 - 69.

- [6] Miao Z L, Shi W Z, Gamba P, et al. An Object-Based Method for Road Network Extraction in VHR Satellite Images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(10): 4853 - 4862.
- [7] Maboudi M, Amini J, Hahn M, et al. Road Network Extraction from VHR Satellite Images Using Context Aware Object Feature Integration and Tensor Voting [J]. *Remote Sensing*, 2016, 8(8): 637.
- [8] 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 Jiguang, Wang Yang, Du Yang, et al. Development and Prospect of Road Extraction Method for Optical Remote Sensing Image [J]. *Journal of Remote Sensing*, 2020, 24(7): 804 - 823.
- [9] 杨晓亮, 文贡坚. 高分辨率遥感影像中提取道路网方法综述[J]. *遥感技术与应用*, 2012, 27(3): 465 - 471.
Yang Xiaoliang, Wen Gongjian. Review on the Methods for Road Network Extraction from High Resolution Remote Sensing Images [J]. *Remote Sensing Technology and Application*, 2012, 27(3): 465 - 471.
- [10] 冯鹏, 高峰. 高分遥感影像中道路信息提取方法综述 [J]. *现代电子技术*, 2015, 38(17): 53 - 57.
Feng Peng, Gao Feng. Method of Road Information Extraction in High Resolution Remote Sensing Images[J]. *Modern Electronics Technique*, 2015, 38(7): 53 - 57.
- [11] Chen Z Y, Deng L A, Luo Y H, et al. Road Extraction in Remote Sensing Data: A Survey[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 112: 102833.
- [12] 李美霖, 芮杰, 金飞, 等. 基于改进 YOLOX 的遥感影像目标检测算法[J]. *吉林大学学报(地球科学版)*, 2023, 53(4): 1313 - 1322.
Li Meilin, Rui Jie, Jin Fei, et al. Remote Sensing Image Target Detection Algorithm Based on Improved YOLOX [J]. *Journal of Jilin University (Earth Science Edition)*, 2023, 53(4): 1313 - 1322.
- [13] 刘霞, 孙英杰. 基于融合残差注意力机制的卷积神经网络地震信号去噪[J]. *吉林大学学报(地球科学版)*, 2023, 53(2): 609 - 621.
Liu Xia, Sun Yingjie. Seismic Signal Denoising Based on Convolutional Neural Network with Residual and Attention Mechanism [J]. *Journal of Jilin University (Earth Science Edition)*, 2023, 53(2): 609 - 621.
- [14] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston: [s. n.], 2015: 3431 - 3440.
- [15] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]// *Proceedings of the Medical Image Computing and Computer Assisted Intervention Society*. Cham: [s. n.], 2015: 234 - 241.
- [16] 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.
- [17] Chen L C, Papandreou G, Kokkinos I, et al. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs[EB/OL]. (2016-06-07) [2023-01-17]. <https://doi.org/10.48550/arXiv.1412.7062>.
- [18] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: SemanticImage Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834 - 848.
- [19] Chen L C, Papandreou G, Schroff F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[EB/OL]. (2017-12-05) [2023-02-12]. <https://doi.org/10.48550/arXiv.1706.05587>.
- [20] Chen L C, Zhu Y, Papandreou G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation [C]// *Proceedings of the European Conference on Computer Vision*. Munich: [s. n.], 2018: 801 - 818.
- [21] Zhang Z X, Liu Q J, Wang Y H. Road Extraction by Deep Residual U-Net [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(5): 749 - 753.
- [22] Wei Y N, Wang Z L, Xu M. Road Structure Refined CNN for Road Extraction in Aerial Image[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(5): 709 - 713.
- [23] 贺浩, 王仕成, 杨东方, 等. 基于 Encoder-Decoder 网络的遥感影像道路提取方法[J]. *测绘学报*, 2019, 48(3): 330 - 338.
He Hao, Wang Shicheng, Yang Dongfang, et al. An Road Extraction Method for Remote Sensing Image Based on Encoder-Decoder Network [J]. *Acta*

- Geodaetical et Cartographical Sinica, 2019, 48(3): 330 - 338.
- [24] 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. Las Vegas: [s. n.], 2016: 770 - 778.
- [25] Sun T, Di Z L, Che P Y, et al. Leveraging Crowdsourced GPS Data for Road Extraction from Aerial Imagery[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: [s. n.], 2019: 7509 - 7518.
- [26] 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.
- [27] Wang C W, Xu R T, Xu S B, et al. DA-Net: Dual Branch Transformer and Adaptive Strip Upsampling for Retinal Vessels Segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Singapore: Springer, 2022: 528 - 538.
- [28] Batra A, Singh S, Pang G, et al. Improved Road Connectivity by Joint Learning of Orientation and Segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: [s. n.], 2019: 10385 - 10393.
- [29] Zhou M T, Sui H G, Chen S X, et al. Large-Scale Road Extraction from High-Resolution Remote Sensing Images based on a Weakly-Supervised Structural and Orientational Consistency Constraint Network[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 193: 234 - 251.
- [30] Ding L, Bruzzone L. DiResNet: Direction-Aware Residual Network for Road Extraction in VHR Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(12): 10243 - 10254.
- [31] Li X G, Wang Y B, Zhang L Q, et al. Topology-Enhanced Urban Road Extraction via a Geographic Feature-Enhanced Network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(12): 8819 - 8830.
- [32] Xu H, He H, Zhang Y, et al. A Comparative Study of Loss Functions for Road Segmentation in Remotely Sensed Road Datasets [J]. International Journal of Applied Earth Observation and Geoinformation, 2023, 116: 103159.
- [33] Mosinska A, Marquez-Neila P, Koziński M, et al. Beyond the Pixel-Wise Loss for Topology-Aware Delineation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: [s. n.], 2018: 3136 - 3145.
- [34] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [EB/OL]. (2015 - 04 - 10) [2023 - 02 - 10]. <https://doi.org/10.48550/arXiv.1409.1556>.
- [35] Oner D, Koziński M, Citraro L, et al. Promoting Connectivity of Network-Like Structures by Enforcing Region Separation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9): 5401 - 5413.
- [36] Guo M H, Lu C Z, Hou Q, et al. SegNeXt: Rethinking Convolutional Attention Design for Semantic Segmentation[EB/OL]. (2022 - 09 - 18) [2023 - 02 - 12]. <https://doi.org/10.48550/arXiv.2209.08575>.
- [37] Gao S H, Cheng M M, Zhao K, et al. Res2net: A New Multi-Scale Backbone Architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2): 652 - 662.
- [38] 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: 1 - 15.
- [39] 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 the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: [s. n.], 2018: 182 - 186.
- [40] Wei Y, Zhang K, Ji S P. Simultaneous Road Surface and Centerline Extraction from Large-Scale Remote Sensing Images Using CNN-Based Segmentation and Tracing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(12): 8919 - 8931.
- [41] 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: 1 - 12.

- [42] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: [s. n.], 2018: 3-19.
- [43] Wei Y, Ji S P. Scribble-Based Weakly Supervised Deep Learning for Road Surface Extraction from Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-12.
- [44] Liu Y H, Yao J, Lu X H, et al. RoadNet: Learning to Comprehensively Analyze Road Networks in Complex Urban Scenes from High-Resolution Remotely Sensed Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 57(4): 2043-2056.
- [45] Lu X Y, Zhong Y F, Zheng Z, et al. Cascaded Multi-Task Road Extraction Network for Road Surface, Centerline, and Edge Extraction [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.
- [46] Chen X, Sun Q, Guo W Y, et al. GA-Net: A Geometry Prior Assisted Neural Network for Road Extraction[J]. International Journal of Applied Earth Observation and Geoinformation, 2022, 114: 103004.
- [47] Bastani F, He S, Abbar S, et al. Roadtracer: Automatic Extraction of Road Networks from Aerial Images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: [s. n.], 2018: 4720-4728.
- [48] Mosinska A, Koziński M, Fua P. Joint Segmentation and Path Classification of Curvilinear Structures[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42(6): 1515-1521.
- [49] Demir I, Koperski K, Lindenbaum D, et al. A Challenge to Parse the Earth Through Satellite Images[EB/OL]. (2018-05-17) [2023-02-20]. <https://doi.org/10.48550/arXiv.1805.06561>.
- [50] Mnih V. Machine Learning for Aerial Image Labeling [D]. Toronto: University of Toronto, 2013.
- [51] Van Etten A, Lindenbaum D, Bacastow T M. Spacenet: A Remote Sensing Dataset and Challenge Series[EB/OL]. (2019-07-15) [2023-02-22]. <https://doi.org/10.48550/arXiv.1807.01232>.
- [52] Zhu Q Q, Zhang Y N, Wang L Z, et al. A Global Context-Aware and Batch-Independent Network for Road Extraction from VHR Satellite Imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 175: 353-365.
- [53] Chen Z Y, Wang C, Li J, et al. Reconstruction Bias U-Net for Road Extraction from Optical Remote Sensing Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 2284-2294.
- [54] Li Y L, Xiang L G, Zhang C L, et al. Fusing Taxi Trajectories and RS Images to Build Road Map via DCNN[J]. IEEE Access, 2019, 7: 161487-161498.
- [55] Grillo A, Krylov V A, Moser G, et al. Road Extraction and Road Width Estimation via Fusion of Aerial Optical Imagery, Geospatial Data, and Street-Level Images [C]//2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Brussels: IEEE, 2021: 2413-2416.
- [56] Liang J, Homayounfar N, Ma W C, et al. Convolutional Recurrent Network for Road Boundary Extraction [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: [s. n.], 2019: 9512-9521.
- [57] Li P L, He X H, Qiao M J, et al. Exploring Multiple Crowdsourced Data to Learn Deep Convolutional Neural Networks for Road Extraction [J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 104: 102544.
- [58] Zhang J, Hu Q W, Li J Y, et al. Learning from GPS Trajectories of Floating Car for CNN-Based Urban Road Extraction with High-Resolution Satellite Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(3): 1836-1847.
- [59] Liang Z Y, Wang T C, Zhang X Y, et al. Tree Energy Loss: Towards Sparsely Annotated Semantic Segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: [s. n.], 2022: 16907-16916.
- [60] Pan Z Y, Jiang P, Wang Y H, et al. Scribble-Supervised Semantic Segmentation by Uncertainty Reduction on Neural Representation and Self-Supervision on Neural Eigenspace [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Nashville: [s. n.], 2021: 7416-7425.