



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 结合 DCNN 与短距条件随机场的遥感影像道路提取
作者: 赫晓慧, 陈明扬, 李盼乐, 田智慧, 周广胜
DOI: 10.13203/j.whugis20210464
收稿日期: 2022-06-24
网络首发日期: 2022-07-27
引用格式: 赫晓慧, 陈明扬, 李盼乐, 田智慧, 周广胜. 结合 DCNN 与短距条件随机场的遥感影像道路提取[J/OL]. 武汉大学学报(信息科学版).
<https://doi.org/10.13203/j.whugis20210464>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20210464

引用格式：

赫晓慧, 陈明扬, 李盼乐, 等. 结合DCNN与短距条件随机场的遥感影像道路提取[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210464(HE Xiaohui, CHEN Mingyang, LI Panle, et al. Road Extraction From Remote Sensing Image by Integrating DCNN With Short Range Conditional Random Field(SRCRF)[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210464)

结合 DCNN 与短距条件随机场的遥感影像道路提取

赫晓慧^{1,3} 陈明扬² 李盼乐² 田智慧^{1,3} 周广胜^{1,3}

1. 郑州大学, 地球科学与技术学院, 河南 郑州, 450052
2. 郑州大学, 信息工程学院, 河南 郑州, 450052
3. 中国气象科学研究院郑州大学生态气象联合实验室 郑州, 450052

摘要：深度卷积神经网络(Deep Convolutional Neural Network, DCNN)在高分辨率遥感图像(High-Resolution Remote Sensing images, HRSIs)自动道路提取领域被广泛应用,但现有方法难以对预测结果中像素间的上下文关系建模。针对此问题,已有研究利用全连接条件随机场(Fully Connected Crf, FullCrf)结合上下文信息对语义分割结果进行二次优化,但无法有效改善道路结构不连续问题。为改善道路结构的完整性,本研究提出一种结合DCNN的短距条件随机场模型(Short Range Conditional Random Filed, SRCRF),SRCRF利用DCNN强大的特征提取能力并控制FullCrf的推理范围缓解过渡平滑现象,解决提道路取结果中的结构不连续、不完整问题。实验结果表明,在Zimbabwe-Roads数据集与Cheng-Roads数据集中SRCRF的F1值相比DCNN分别上升大约4.01%与3.73%,相比FullCrf分别上升大约3.25%和2.28%。

关键词：遥感影像; 道路提取; 深度学习; 条件随机场; 语义分割;

中图分类号：TP75 **文献标识码：**A

1 引言

道路提取在遥感领域一直是热点研究问题^[1],它是许多应用的必要步骤,如城市规划^[2]、智能交通^[3]、地理信息系统更新^[4]等。HRSIs因具有丰富的地理信息与细节信息是道路提取任务很好的数据源。在实际应用中依靠手工标记从HRSIs中提取道路十分费时费力,并且道路不规则与复杂的结构容易导

致错误标记,因此自动道路提取变得越来越重要。

道路提取实质上是语义分割问题。近年来,研究者提出一系列机器学习算法进行语义分割,这些算法包括随机森林^[5],支持向量机(Support Vector Machine,SVM)^[6],深度卷积神经网络DCNN^[7]等。DCNN在提取局部特征和利用较小的感受野进行预

收稿日期: 2022-06-24

项目资助: 第二次青藏高原综合考察研究项目(2019QZKK0106)

第一作者: 赫晓慧, 教授, 博士生导师, 主要从事人工智能, 遥感影像处理及数据挖掘。hexh@zzu.edu.cn

测方面非常强大,因此在语义分割领域被认为是一个高效的方法。在遥感领域,DCNN被广泛应用于[8-11]。Abderrahim等[12]利用UNet[13]从复杂环境的遥感影像中进行道路提取。Cheng等[14]提出一个新的端到端级联网络CasNet,其由两个子网络级联组成,后部的网络可以充分利用前部网络提取到的特征从而获取更好的道路提取结果。Zhang等[15]提出一个新的模型使UNet[13]与ResNet[16]结合从而减少训练所需的参数数量,使网络更加容易训练并使道路提取结果更准确。He等[17]提出将空间金字塔池化结构[18]与UNet网络结合以提高细节特征的提取能力进行遥感影像道路提取。宋等[19]提出在编码阶段引入注意力机制以筛选出重要特征,并在卷积层后加入空间金字塔池化模块以获取不同尺度的道路特征。

但是本文在使用以上实验[8,13,20]过程中发现在道路提取结果中会出现道路不连续与道路边缘细节提取不完整等问题,特别在建筑物阴影或者树木遮挡区域。其中主要原因是1)遥感影像中道路区域存在树木,汽车等小型物体的遮挡,传统DCNN方法难以识别,造成道路识别出现断点等问题[21]。2)UNet,PspNet,SegNet三种神经网络均采用多层卷积进行特征提取,特征提取过程的编码阶段使得信息压缩导致部分细节信息丢失,并且以上三种方法缺少利用预测结果中语义空间上下文关系的能力[14]使得预测结果中道路提取不够完整。在这种情况下,研究人员开始探索引入图模型如条件随机场(Conditional Random Field, CRF)对语义分割结果进一步优化。CRF有能力同时对观测图像与预测结果之间的上下文进行建模。其中具有代表性的例子是采用由Philipp等[22]提出具有高效估计推理算法的FullCrf。FullCrf在图像的所有像素间构建成对势函数形成像素级图模型,这使其在给每个像素分配标签时考虑到像素间的空间相关性与光谱相关性。并且FullCrf推理算法将对势函数定义为高斯核函数的线性组合降低了推理的时间复杂度。FullCrf近年来已经成为图像二次优化的标准工具[23]。针对道路检测问题Rao等[21]利用FullCrf优化FCN输出的道路分类结果。Alam等[24]在高光谱影

像中使用聚类算法产生超像素,随后将卷积神经网络(Convolutional Neural Network, CNN)与CRF结合以利用两者的优势以提取出更精确的目标。左等[25]引入可变型卷积层以增强网络对空间变换的适应能力,并使用FullCrf对提取网络结果进行二次优化。

这些方法虽然利用空间相关性与光谱相关性两种上下文信息对预测之间的交互进行建模,但并不适用于高分辨率遥感影像道路提取任务。在实验中FullCrf容易对道路过渡平滑,导致道路细节缺失从而不能有效解决道路不连续问题。主要原因是高分辨率遥感影像中道路形态具有细长及影像占比小等特点,FullCrf构建的数学模型中特征向量(道路-背景)的占比大于特征向量(道路-道路)的占比,使得FullCrf在推理过程中过于依赖特征向量(道路-背景),增大了部分道路像素点被推理为背景像素点的概率,从而引起道路的过渡平滑。

由于FullCrf模型中特征向量维度高导致模型复杂度也较高,Teichmann等[26]实验证明FullCrf推理要比DCNN推理慢两个数量级。这在一定程度上阻碍FullCrf作为高分辨率遥感影像目标提取等任务的后处理方法的深入研究和实验。考虑到以上问题,并受到ConvCrf[26]的启发,本文提出一种新颖的短距条件随机场(SRCRF)。其在推理过程中将特征向量的构造范围由全连接结构优化为K-邻域结构,从而提高特征向量(道路-道路)的占比,使得SRCRF相比于FullCrf更适用与高分辨率遥感影像道路任务,对道路提取结果的优化更加完善。

文章其余部分组织如下。第一章节讨论与文章有关的相关方法。第二章节详细介绍SRCRF原理。第三章节展示与分析实验结果。第四章节总结全文,以及今后的研究方向。

2 短距条件随机场

高分辨率遥感影像有丰富的地理信息,是道路提取很好的数据源,但DCNN独立地对每个像素进行预测,忽视了相邻像素间的关系,在树木与建筑物遮挡区域的预测结果中出现道路不连续现象。为解决道路提取任务中对预测结果的二次优化以及缓解FullCrf对道路的过渡平滑问题,本文提出针对道路提取的SRCRF模型。在SRCRF模型中本文根据

遥感影像道路结构定义了一元势函数与二元势函数来考虑上下文关系，并且提出 K 邻域推理算法用以结合光谱-空间上下文信息推理出观测像素对应的最佳标签。针对高分辨遥感影像的道路提取模型 SRCRF 具体如图 1 所示。

为清楚描述 SRCRF 算法，本文将包含 N 个像素 k 个类别的遥感影像定义为 I ， I 的分割结果建模为 $X = \{X_1, \dots, X_n\}$ 。每一个变量 X_i 可以取值 $\{1, \dots, k\}$ 。语义分割问题可以看作是找到使后验概率 $P(X | I)$ 最大化的图像标签集，最大化后验概率 $P(X | I)$ 等价于最小化 Gibbs 能量函数。

本文提出的 SRCRF 模型是一种判别分类方法，它直接以观察到的图像与标签 x 的后验概率进行建模为 Gibbs 能量函数，形式如下：

$$P(X = x | I) = \frac{1}{Z(I)} \exp(-E(x | I)) \quad (1)$$

其中 $Z(I)$ 是配分函数， $E(X)$ 是能量函数，在

势团中定义该函数以对随机变量之间的关系进行建模。根据各种类型的势团，势函数可以分为一元势函数，二元势函数与高阶势函数。对于遥感影像道路提取任务来说，具有一元和二元势函数的 CRF 模型被广泛使用^[21,27]。因此 SRCRF 的构造也采用一元势函数与二元势函数组合的方式。SRCRF 相应的能量函数定义为：

$$E(x_i | I) = \sum_{d_x \leq k} \varphi_u(x_i, d_x) + \sum_{d_x, d_y \leq k} \varphi_p(x_i, x_j, d_x, d_y) \quad (2)$$

式中， $\varphi_u()$ 是一元势函数， $\varphi_p()$ 是二元势函数， d_x, d_y 分别代表在 X 中距 x_i 和 x_j 的距离， K 代表 SRCRF 中设定的曼哈顿距离。

道路提取任务实际上是像素分类问题，分类问题旨在基于贝叶斯最大后验（MAP）找到使后验概率 $P(X | I)$ 最大化的状态序列 X 。SRCRF 模型的图像分类问题的最大后验概率 X_{MAP} 由下式给出：

$$X_{MAP} = \arg \max P(X | I) = \arg \min E(x | I) \quad (3)$$

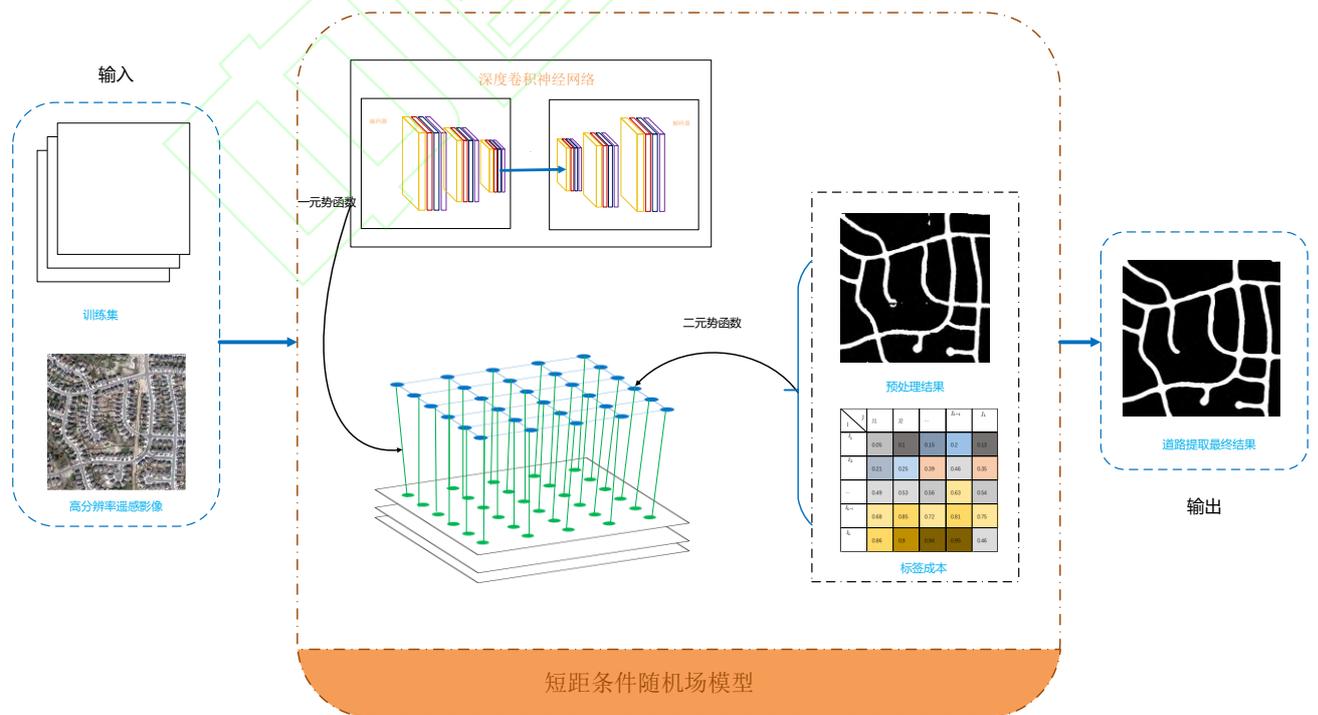


图 1 SRCRF 模型整体流程图

Fig. 1 overall flow chart of SRCRF model

2.1 基于道路预分割的一元势函数

一元势函数 $\varphi_u(\bullet)$ 是 CRF 的重要组成部分，负责对观测像素与标签之间的关系进行建模，可以看作观测像素对应的类成员概率。FullCrf 中类成员概率通常由传统判别模型 SVM, CNN 等根据像素光谱特征计算得出。SRCRF 采用判别模型 DCNN 作为分类器。相比于传统的判别模型，DCNN 具有更强大的特征提取能力，这种能力在道路背景复杂的高分辨率遥感影像中十分关键。由 DCNN 定义的一元势函数形式如下所示：

$$\varphi_u(x_i) = -\log P(x_i) \quad (4)$$

其中 $P(\bullet)$ 代表 DCNN 的输出结果。

在推理过程中将一元势函数最小化的过程等同于图像的非上下文贝叶斯分类。当像素被 DCNN 正确分类的时候，一元势函数的能量比较低，相反当像素被错误分类时一元势函数项能量会变高即 $\varphi_u(\bullet)$ 值变大。

2.2 基于光谱特征与空间特征的二元势函数

空间先验知识符合人类的认知，在分类中对于确定像素的类型非常重要。二元势函数通过构造特征向量，对周边节点的依赖性建模从而将道路空间先验知识考虑在内，并且二元势函数 $\varphi_p(\bullet)$ 利用空间先验知识使分类结果具有局部一致性，使得光谱特征相同的邻像素在输出分类图中具有相同的标签。FullCrf 与 SRCRF 中由二元势函数构造特征向量的过程如图 2 所示。

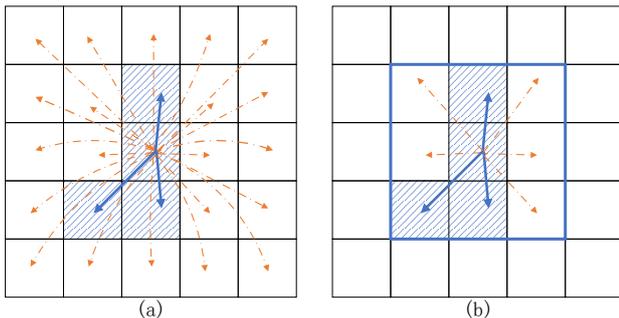


图 2 FullCrf (a) 特征向量与 SRCRF (b) 特征向量构造示意图

Fig. 2 Schematic diagram of FullCrf (a) and SRCRF (b) feature vector construction

图 2 中网格代表遥感影像，蓝色网状区域代表道路，空白区域代表背景。蓝色方框代表 K 为 1 时 SRCRF 的特征向量构造范围。橙色箭头为特征向量（道路-背景），蓝色箭头为特征向量（道路-道路）。从 (a) 可以看出 FullCrf 构造特征向量时特征向量（道路-背景）数量多于特征向量（道路-道路）的数量，约为特征向量（道路-道路）的 7 倍，而从 (b) 可以看出，SRCRF 在 K 为 1 的范围中构造的特征向量（道路-背景）数量仅为特征向量（道路-道路）的 1.7 倍。FullCrf 中两种特征向量占比相对于 SRCRF 更不平衡，使得 FullCrf 在推理过程中过于依赖占比大的特征向量即特征向量（道路-背景），增大了部分道路像素点被推理为背景的概率，从而引起道路的过渡平滑。而 SRCRF 通过缓解 HRSIs 道路提取任务中特征向量的占比不平衡达到了对道路更好的二次优化效果。

Krhenbühl 等 [28] 提出使用高斯滤波的方式进行推理可以减小其计算复杂度，为使分类结果具有局部一致性缓解道路提取结果中的道路断点问题以及提高推断速度，在 SRCRF 中本文将二元势函数定义为高斯核函数的线性组合，具体如下所示：

$$\varphi_p(x_i, x_j, d_x, d_y) = \mu(x_i, x_j, d_x, d_y) \sum_{m=1}^M \omega^{(m)} k_G^{(m)}(f_i, f_j, d_x, d_y) \quad (5)$$

$k_G^{(m)}(f_i, f_j, d_x, d_y)$ 是边缘特征函数用来度量高分辨率遥感影像相邻像素间的外观差异，其中包括两个手工设计的高斯核函数， f_i 是像素 i 的特征

向量。 $k_G^{(m)}(f_i, f_j, d_x, d_y)$ 的具体定义如下：

$$k_G^{(m)}(f_i, f_j, d_x, d_y) = k_\alpha(f_{i+d_x+d_y}^I, f_{j+d_x+d_y}^I) + k_\beta(f_{i+d_x+d_y}^I, f_{j+d_x+d_y}^I) \quad (6)$$

$$k_\alpha(f_{i+d_x+d_y}^I, f_{j+d_x+d_y}^I) = \omega^{(1)} \exp\left(-\frac{p_{i+d_x+d_y} - p_{j+d_x+d_y}}{2\theta_\alpha^2} - \frac{l_{i+d_x+d_y} - l_{j+d_x+d_y}}{2\theta_\beta^2}\right) \quad (7)$$

$$k_{\beta}(f_{i+d_x+d_y}^I, f_{j+d_x+d_y}^I) = \omega^{(2)} \exp\left(-\frac{|I_{i+d_x+d_y} - I_{j+d_x+d_y}|}{2\theta_{\gamma}^2}\right) \quad (8)$$

$\mu(x_i, x_j)$ 表示兼容性函数, 一个常用的兼容性函数[3]是 Potts 模型:

$$\mu(x_i, x_j) = |x_i \neq x_j| \quad (9)$$

其中 $k_{\alpha}(\bullet)$ 为使用颜色特征 I_i, I_j 的高斯核, 其促进相同颜色的像素具有相同标签。 $k_{\beta}(\bullet)$ 是使用空间位置坐标特征 P_i, P_j 的高斯核, 可以去除小的孤立区域。参数 $\omega^{(m)}$ 为确定平滑度的常数。

2.3 K 邻域平均场推理算法

SRCRF 结合道路先验概率与空间、光谱特征对道路提取结果进行优化, 优化过程可以归纳为图像的二次标注问题, 在条件随机场中准确地图像标注是个 NP-hard 问题。同时高分辨率遥感影像中道路形态往往具有狭窄, 结构复杂的特征, FullCrf 构建的数学模型中特征向量(道路-背景)的数量大于特征向量(道路-道路)的数量, 使得在 FullCrf 模型中影像占比较小的目标容易被误标记为占比较大的目标。考虑到这些问题, 本文提出 K 邻域平均场推理算法, 以估算出每个像素对应的最佳标签。该算法以卷积的方式将条件随机场的推理范围控制在在半径为 K 的范围内以提高特征向量(道路-道路)的占比, 从而充分利用道路的空间上下文信息与光谱特征上下文信息。并基于平均场近似理论^[29]估算出每个像素对应的最佳标签。这使得模型总体的特征向量维度减少的同时又在一定程度上提高了特征向量(道路-道路)的占比, 从而解决 FullCrf 过度平滑导致的细节缺失无法有效改善道路形态的问题。半径为 K 的 SRCRF 上下文消息传递流程具体如图 3 所示。

如图 3 所示 K 取 1 时 SRCRF 的边长为 3, 每一个节点对应图像中的一个像素, 点与点之间的边代表像素间的关系, 红色的点代表中心像素点, 蓝色

的代表相邻像素。(i,u)表示像素点 i 的标签为 u。 $\mu(x_i, x_j)$ 代表(i,u), (j,v)同时将 i 和 j 在各自位置上赋予标签 u 和 v 的成本。上下文消息传递图 3(a)所示, 在 i 点 3×3 范围内执行具有边缘特征核的卷积操作, 得出 i 点的能量数值 $Q_i^v(I)$ 。兼容性转换如 3(b)所示, $\mu(x_i, x_j)$ 与 $Q_i^v(I)$ 进行点乘。完整的 SRCRF 特征降维推理算法如表 1 所示。

3 实验结果与分析

这一章节首先描述实验设置, 其次介绍所用实验数据集与评价指标, 最后展示实验的结果。

3.1 实验设置

实验算法的架构为 Tensorflow, CUDA Version 10.0.130。网络的参数更新算法为 SGD 梯度下降算法, 学习率设置为 $2e-7$, batchsize 为 1, 程序运行平台的 GPU 为 NVIDIA Tesla P100, RAM 的大小为 16GB, CPU 为 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz。

3.1.1 实验数据集

实验采用两个数据集: 1) Zimbabwe-Roads 数据集; 2) Cheng-Roads 数据集。

Zimbabwe-Roads 数据集是从 Bing Maps 收集建立, 对应的标签从 OpenStreetMap (OSM) 产生。

Zimbabwe-Roads 数据集的空间分辨率是 $0.28 m^2$ 。经过手工完善标签, 筛选等步骤后, 数据集包含 248 张尺寸为 1500×1500 的遥感影像。其中 197 张用于训练, 10 张用于验证, 41 张用于测试。

Cheng-Roads 数据集为公开数据集, 公布在[20]。包含 224 张从谷歌地球下载的尺寸为 640×640 的遥感影像。这大约覆盖 47341km 的道路, 并且道路的标签都是手工标注。每个像素的空间分辨率为 $1.2 m^2$ 。在这之中随机挑选 180 张用来训练, 14 张用来验证, 30 张用来测试模型。

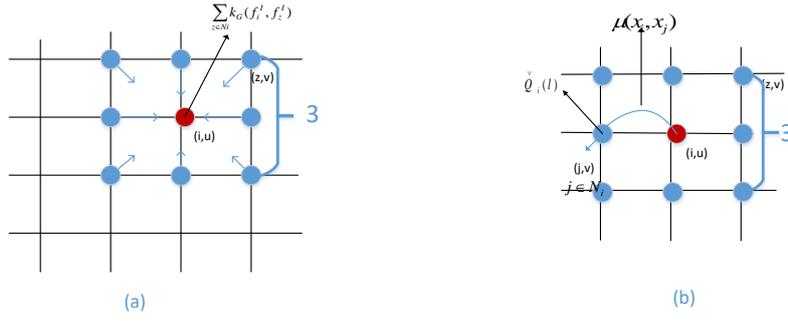


图3 SRCRF 上下文消息传递 (a) 与兼容性转换 (b) 示意图

Fig.3 schematic diagram of SRCRF context messaging (a) and compatibility transformation (b)

表 1 K 邻域平均场推理算法

Tab.1 K-Neighborhood mean field reasoning algorithm

K 邻域平均场推理算法	
1: Initialize:	
	$Q_i = \frac{1}{Z_i} \exp(-\varphi_u(x_i I))$
2: While not converged do:	
3: $Q_i^m(l) \leftarrow \sum_{d_x, d_y \leq k} k_G^{(m)}(f_i, f_j, d_x, d_y) Q_i$	▷ 上下文消息传递
4: $Q_i^v(l) \leftarrow \sum_m \omega^{(m)} Q_i^m(l)$	▷ 加权输出
5	: ▷ 兼容性转换
$Q_i^{\wedge}(l) \leftarrow \sum_{d_x, d_y \leq k} \mu(x_{i+d_x+d_y}, x_{j+d_x+d_y}) Q_i^v(l)$	
6: $Q_i(l) \leftarrow \varphi_u(x_i) + Q_i^{\wedge}(l)$	▷ 加上一元势函数
7: $Q_i(l) = \frac{1}{Z_i} \exp(-Q_i(l))$	▷ 标准化 eg. softmax
8: End while	

两个数据集集中的大多数图片都具有复杂的背景以及一些由树木与汽车等物体造成的遮挡。所有这些因素使获得令人满意的遥感图像结果非常具有挑战性。

3.1.2 分割性能评价指标

文章采用召回率 (R)、准确率 (P) 和 F1 值作为遥感图像分割结果的评价指标，计算公式如下所示。

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

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

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (12)$$

TP (True Positive) 代表将道路预测为道路的像素数量。FN (False Negative) 代表将道路预测为背景的像素数量。FP (False Positive) 代表将背景预测为道路的像素数量。TN (True Negative) 代表将背景预

测为背景的像素数量。准确率 P 代表道路的分割结果与标签的交集占标签的比例，召回率 R 代表预测出的正确道路占所有预测道路中的比值。F1 值为结合以上两种评价指标的特点，对分割结果进行综合评价。

3.2 短距条件随机场与其他方法的对比

为证明 SRCRF 的有效性，本文在 Zimbabwe-Roads 数据集与 Cheng-Roads 数据集上进行实验。为保证实验的客观性，除 SRCRF 的超参数 K 以外其余超参数与 FullCrf 保持一致。Zimbabwe-Roads 测试集的结果如表 2 所示，Cheng-Roads 测试集结果如表 3 所示。

表 2, 表 3 提供 SRCRF 与其他方法在 Zimbabwe-Roads 测试集与 Cheng-Roads 测试集上定量结果的对比。左边第一列是模型的名称，其余从左往右是三种评价指标 Precision, Recall, F1。在 Zimbabwe-Roads 数据集中由 UNet+FullCrf 方法取得的 F1 值达到 0.8014，超过实验选取的三种深度学习方法至少 0.76%。由 UNet+SRCRF 取得的 F1 值达到 0.8339，这与 UNet+FullCrf 方法相比高 3.25%。UNet 是目标提取的常用算法，在 Zimbabwe-Roads 数据集中它的 F1 值只有 0.7938 低于 UNet+SRCRF 的结果。在 Cheng-Roads 数据集中 UNet+FullCrf 方法的 F1 值与实验中道路提取结果最好的深度学习方法 UNet 相比提升 1.45%。而本文提出的 UNet+SRCRF 方法取得 F1 值为 0.8991 相比 UNet+FullCrf 方法提高 2.28%。

表 2 SRCRF 与其他方法的在 Zimbabwe-Roads 数据集上的结果比较

Tab.2 Comparison of the results of the proposed method with other methods on Zimbabwe-Roads

模型	准确率 (P)	召回率 (R)	F1
PspNet	0.7720	0.7948	0.7832
SegNet	0.7655	0.7616	0.7635
UNet	0.8363	0.7555	0.7938
UNet+FullCrf	0.8133	0.7898	0.8014
UNet+SRCRF	0.8403	0.8275	0.8339

表 3 SRCRF 与其他方法的在 Cheng-Roads 数据集上的结果比较

Tab.3 Comparison of the results of the proposed method

with other methods on Cheng-Roads

模型	准确率 (P)	召回率 (R)	F1
PspNet	0.8689	0.7641	0.8131
SegNet	0.8386	0.8273	0.8329
UNet	0.8643	0.8594	0.8618
UNet+FullCrf	0.8834	0.8694	0.8763
UNet+SRCRF	0.9160	0.8829	0.8991

PspNet, SegNet, UNet 在 Zimbabwe-Roads 数据集与 Cheng-Roads 数据集的可视化道路提取结果如图 4 所示。图 4 中展示了四幅遥感影像，这四幅影像具有乡村与城市街区两种背景，以及沙土路与柏油路两种道路性质，并且道路形态复杂各异。另外我们将影像中道路受到遮挡的部分区域用黄色方框标记并放大如 (b) 列所示，以更清楚看出树木等物体的遮挡现象。并将三种方法输出结果的对应区域也用黄色方框标记如 (d) (e) (f) 列所示。

由图 4 可以看出在两种高分辨率遥感影像数据集中，三种方法的道路提取结果都存在部分道路细节缺失，道路中断等现象，尤其在树木等小型物体遮挡区域。并且可以看出图 4 中 (f) 列结果略优于 (d) 列与 (e) 列其他两种方法。为更好展示文章方法的优势，我们采用三种方法中 F1 值最高并且可视化效果最好的 UNet 作为 FullCrf, SRCRF 的对比方法。并把 UNet 结果作为 FullCrf 与 SRCRF 后处理优化的基础数据。

UNet, FullCrf 与 SRCRF 在 Zimbabwe-Roads 数据集与 Cheng-Roads 数据集的可视化道路提取结果如图 5 所示。图 5 在图 4 的基础上增加了一幅具有较大空地的乡村遥感影像以及一幅具有城市高速公路背景的遥感影像。

从图 5 中 (c) 列可以看出 UNet 网络的语义分割结果可以看出路网结构络已被很好地识别，但道路边缘存在毛刺以及存在一些远离道路中心线的噪声点和由建筑阴影或树木遮挡造成的道路分割断点。经 FullCrf 处理，边缘毛刺、噪声和断点问题已有改善，但一些较大噪声与较宽的断点仍没有被完全消除，而 SRCRF 在边界清晰与连续性两个方面相比于 FullCrf 都取得了更好结果。这是因为对于 HRSIs 道路提取任务来说 FullCrf 的推理范围过大使得 FullCrf 中特征向量 (道路-背景) 的数量也远大于特

征向量（道路-道路）的数量，导致道路被误认为背景。

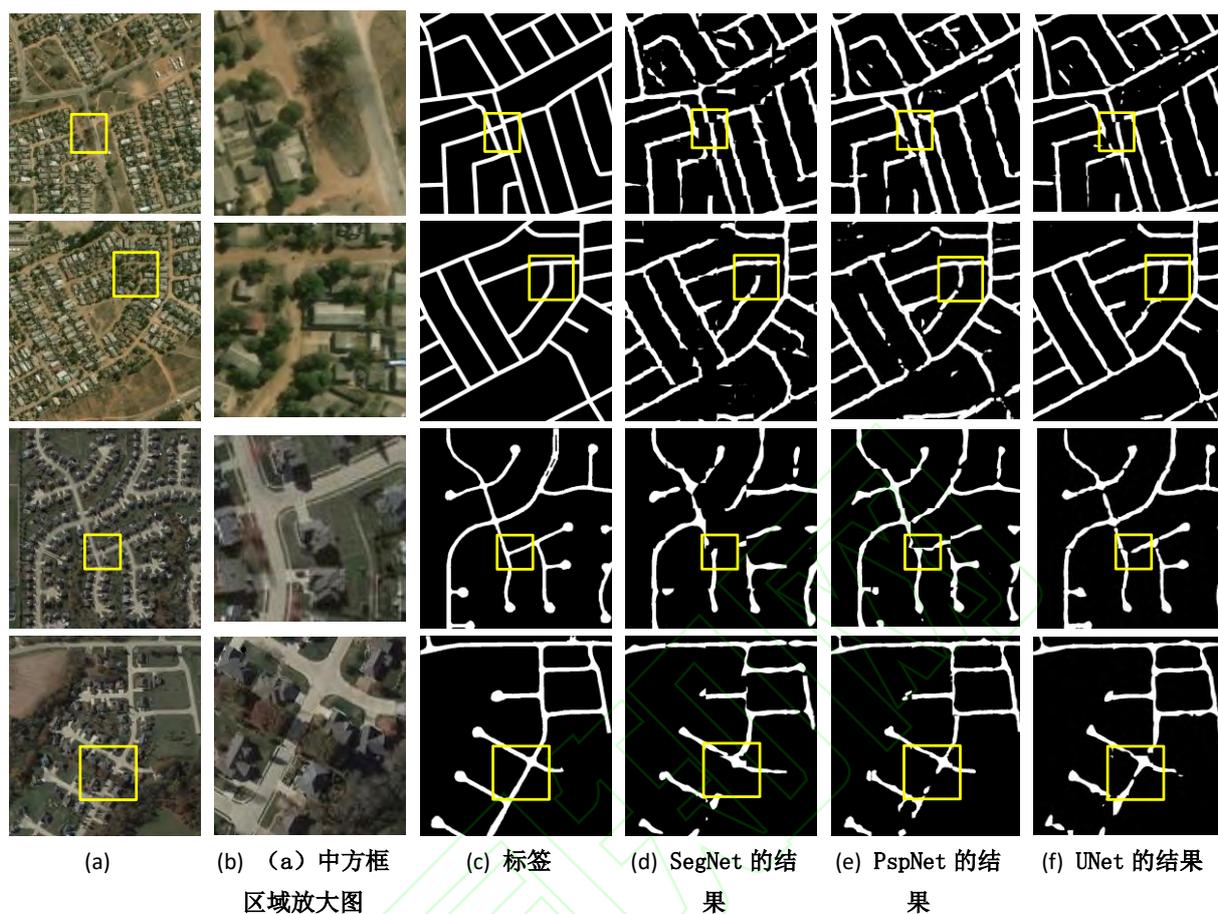
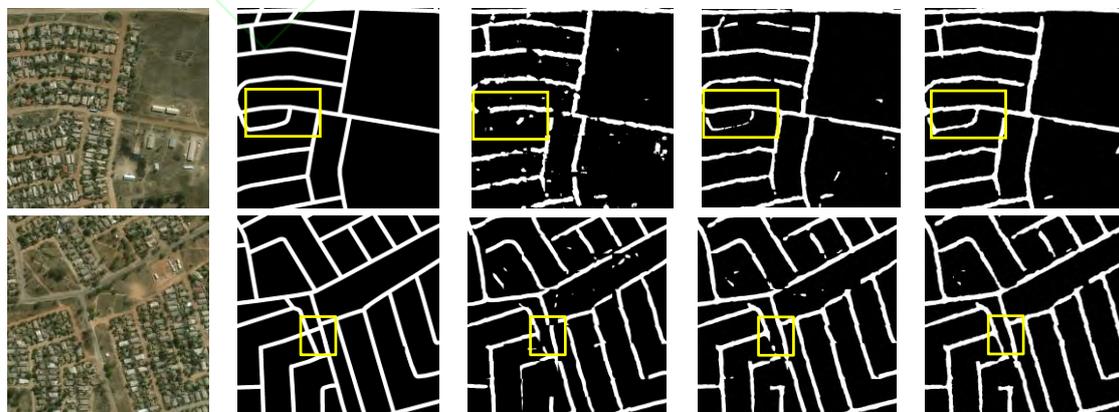


图 4 在 Cheng-Roads 与 Zimbabwe-Roads 数据集中 SegNet、PspNet 与 UNet 的可视化结果对比。前两行为 Zimbabwe-Roads 数据集，后两行为 Cheng-Roads 数据集

Fig.4 Comparison of visualization results of segnet, pspnet and UNET in Cheng roads and Zimbabwe roads data sets. The first two lines are Zimbabwe roads dataset, and the last two lines are Cheng roads dataset



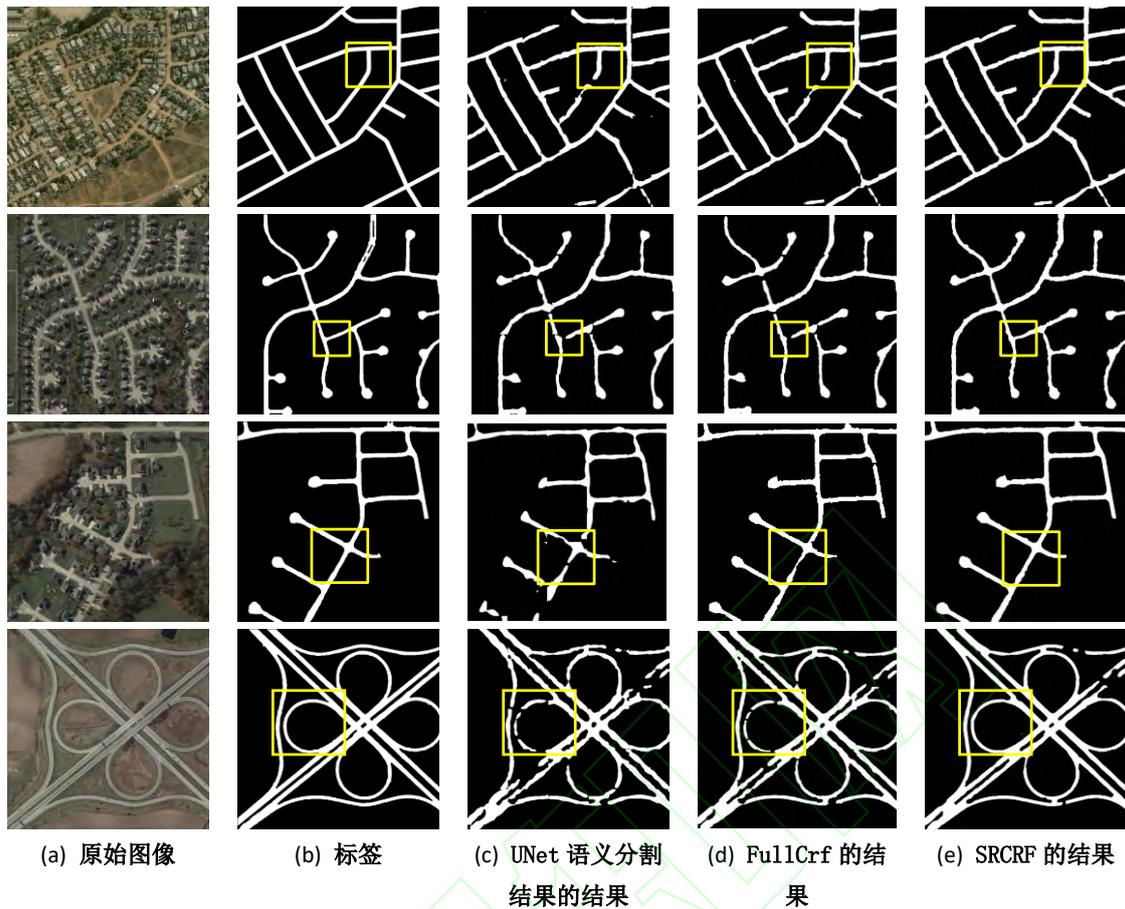


图 5 在 Cheng-Roads 与 Zimbabwe-Roads 数据集中 SRCRF 与其他方法的可视化结果对比。前三行为 Zimbabwe-Roads 数据集，后三行为 Cheng-Roads 数据集

Fig.5 Comparison of the visualization results of SRCRF and other methods in Cheng-roads and Zimbabwe-roads datasets. The first three rows are Zimbabwe-roads, and the last three rows are Cheng-roads

3.3 短距条件随机场计算性能评价

为更好研究短距条件随机场的时间复杂度，本文在 Zimbabwe-Roads 和 Cheng-Roads 数据集上分别做 6 组实验。分别包括 UNet 在 Zimbabwe-Roads 和 Cheng-Roads 数据集上提取一张遥感影像中道路的耗时于 FullCrf 后处理一张图片的耗时，以及不同作用范围的短距条件随机场耗时，具体实验结果如表 4 所示。

表 4 为 FullCrf 与 K 为 1、3、5、7、9 的 SRCRF 的用时对比。表格有 3 行 7 列，第一行是各个测试模型名字，第二列是 FullCrf 后处理一张遥感图像的用时，第 3 列至第 7 列是短距条件随机场在不同作用范围下处理一张遥感图像的用时。

根据表 4 内容可以看出当 SRCRF 的 K 为 1 时，处理一张图片的耗时相比全连接条件随机场减少两个数量级。虽然随着 SRCRF 作用范围的增大推理

时间也逐步增加，但总体来说仍比 FullCrf 的耗时减少一个数量级。

3.4 距离超参数 K 的敏感性分析

不同的高分辨率遥感影像数据集的道路背景占比不同，因此对应的最佳推理范围 K 也不相同。为验证超参数 K 对道路提取结果的影响，本文在 Cheng-Roads 和 zimbabwe-Roads 数据集上进行对比实验。其中 Cheng-Roads 数据集中道路占比 21.261%，zimbabwe-Roads 数据集中道路占比 16.164%，K 取值从 1 到 9，间隔为 2。实验的定量结果如表 5 所示。

表 5 为具有不同作用范围 SRCRF 之间的定量比较。origin 列是 UNet 的定量分割结果。1、3、5、7、9 列为不同推理范围的 SRCRF 对 UNet 结果定量优化结果。

从表 5 可以看出随着 K 值的增大 F1 值呈现先上升后下降的趋势。在 Zimbabwe-Roads 数据集中 SRCRF 的 K 值为 1 时，道路分割结果的 F1 值最高，

提升为 4.01%。在 Cheng-Roads 数据集中 K 值为 3 时分割效果最好，道路 F1 值提升为 2.67%。实验表明当影像分辨率与道路背景占比不同时，SRCRF 对应的最优 K 值也不相同。这是因为当道路背景占比小时，SRCRF 中特征向量（道路的道路）与特征向量（道路-背景）的比例也小，此时需要较小的 K 值避免对道路的过渡平滑。当道路占比大时，

SRCRF 中特征向量（道路-道路）与特征向量（道路-背景）的比例也会增加。此时过渡平滑问题对道路二次优化的影响减小。SRCRF 则可以通过 K 值的增大以考虑更多的上下文信息对道路进行更较准确优化从而提高道路的 F1 值。

4 结论

相比于使用传统深度学习方法进行高分辨率遥感影像道路提取，本文提出一种新的遥感影像道路提取方案 Short Range Conditional Random Filed, SRCRF。这个方案结合 DCNN 的优势并将传统条件随机场的全连接结构优化为 K 邻域结构，缩小了推理范围，提高了特征向量（道路-道路）的占比。这

使 SRCRF 相比与 FullCrf 能更好利用影像颜色特征与空间特征对深度学习输出的道路提取结果进行精确优化。根据 SRCRF 在 Cheng-Roads 与 Zimbabwe-Roads 数据集上的结果显示，该方法较 DCNN 与 FullCrf 均有性能提升，并且较 FullCrf 用时缩短一个数量级。这项工作也为以后的研究提供了理论基础和应用基础。在未来的工作中，我们将进一步研究学习高斯特征的潜力，研究更复杂的 CRF 架构，以更好地捕获全局上下文信息。最后，我们特别感兴趣在其他领域探索 SRCRF 的应用潜力，例如，建筑物提取、汽车提取、湖泊提取等领域。

表 4 SRCRF 在 Zimbabwe-Roads and Cheng-Roads 于其他方法的计算耗时对比。

Tab .4 Comparison of the Calculation time of the proposed method with other methods on Zimbabwe-Roads and Cheng-Roads.

数据集	FullCrf	SRCRF1	SRCRF3	SRCRF5	SRCRF7	SRCRF9
Zimbabwe	557.93ms	5.2537ms	5.3459ms	8.8414ms	13.5821ms	19.0186ms
Cheng	139.02ms	2.1306ms	3.5228ms	5.1925ms	8.3709ms	13.0393ms

表 5 SRCRF 在不同推理范围下的定量比较

Tab .5 Quantitative comparison of SRCRF in different inference scopes

数据集	SRCRF 推理范围																	
	预分类结果			1			3			5			7			9		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Zimbabwe	0.8363	0.7555	0.7938	0.8403	0.8275	0.8339	0.8375	0.7973	0.8169	0.8300	0.7805	0.8045	0.8247	0.7780	0.8007	0.8138	0.7664	0.7893
Cheng	0.8646	0.8594	0.8618	0.8751	0.8697	0.8724	0.9160	0.8829	0.8991	0.9016	0.8905	0.8960	0.8789	0.8785	0.8787	0.8697	0.8701	0.8699

参考文献(References):

- [1] Abdollahi A, Pradhan B, Shukla N, et al. Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-of-the-Art Review[J]. *Remote Sensing*, 2020, 12(9): 1444
- [2] Soulard C E, Acevedo W, Stehman S V. Removing Rural Roads from the National Land Cover Database to Create Improved Urban Maps for the United States, 1992 to 2011[J]. *Photogrammetric Engineering & Remote Sensing*, 2018, 84(2): 101-109
- [3] Li D M, Deng L B, Cai Z M, et al. Notice of Retraction: Intelligent Transportation System in Macao Based on Deep Self-Coding Learning[J]. *IEEE Transactions on Industrial Informatics*, 2018, 14(7): 3253-3260
- [4] Vargas-Muñoz J E, Lobry S, Falcão A X, et al. Correcting Rural Building Annotations in OpenStreetMap Using Convolutional Neural Networks[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 147: 283-293
- [5] Xu Y Y, Xie Z, Wu L, et al. Multilane Roads Extracted from the OpenStreetMap Urban Road Network Using Random Forests[J]. *Transactions in GIS*, 2019, 23(2): 224-240
- [6] Bruzzone L, Persello C. A Novel Context-Sensitive Semisupervised SVM Classifier Robust to Mislabeled Training Samples[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(7): 2142-2154
- [7] Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651
- [8] 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
- [9] Ge Yun, Jiang Shunliang, Ye Famao, et al. Remote Sensing Image Retrieval Using Pre-Trained Convolutional Neural Networks Based on ImageNet[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(1): 67-73 (葛芸, 江顺亮, 叶发茂, 等. 基于 ImageNet 预训练卷积神经网络的遥感图像检索[J]. *武汉大学学报·信息科学版*, 2018, 43(1): 67-73)
- [10] Men Jilin, Liu Yueyan, Zhang Bin, et al. Land Use Classification Based on Multi-Structure Convolution Neural Network Features Cascading[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(12): 1841-1848 (门计林, 刘越岩, 张斌, 等. 多结构卷积神经网络特征级联的高分影像土地利用分类[J]. *武汉大学学报·信息科学版*, 2019, 44(12): 1841-1848)
- [11] He Hao, Wang Shicheng, Yang Dongdang, et al. An road extraction for remote sensing image based on Encoder-Decoder network[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(3): 330-338. (贺浩, 王仕成, 杨东方, et al. 基于 Encoder-Decoder 网络的遥感影像道路提取方法[J]. *测绘学报*, 2019, 048(003): 330-338.)
- [12] Abderrahim N Y Q, Abderrahim S, Rida A. Road Segmentation using U-Net architecture[C]//2020 IEEE International conference of Moroccan Geomatics (Morgeo). Casablanca, Morocco.: 1-4
- [13] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241
- [14] Cheng G L, Wang Y, Xu S B, et al. Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(6): 3322-3337
- [15] 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
- [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA.: 770-778
- [17] He H, Yang D F, Wang S C, et al. Road Extraction by

- Using Atrous Spatial Pyramid Pooling Integrated Encoder-Decoder Network and Structural Similarity Loss[J]. *Remote Sensing*, 2019, 11(9): 1015
- [18] Chen L C, Papandreou G, Schroff F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[EB/OL]. 2017: arXiv: 1706.05587. <https://arxiv.org/abs/1706.05587>
- [19] Song Tingqiang, Liu Tongxin, Zong Da, et al. Research on Road Extraction Method from Remote Sensing Images Based on Improved U-Net Network[J]. *Computer Engineering and Applications*, 2021, 57(14): 209-216 (宋廷强, 刘童心, 宗达, 等. 改进 U-Net 网络的遥感影像道路提取方法研究[J]. *计算机工程与应用*, 2021, 57(14): 209-216)
- [20] 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. Honolulu, HI, USA.: 6230-6239
- [21] Rao Y B, Liu W, Pu J S, et al. Roads detection of aerial image with FCN-CRF model[C]//2018 IEEE Visual Communications and Image Processing. Taichung, China.: 1-4
- [22] Krähenbühl P, Koltun V. Efficient Inference in Fully Connected CRFS with Gaussian Edge Potentials[EB/OL]. 2012: arXiv: 1210.5644. <https://arxiv.org/abs/1210.5644>
- [23] Alam F I, Zhou J, Liew A W C, et al. Conditional Random Field and Deep Feature Learning for Hyperspectral Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(3): 1612-1628
- [24] Alam F I, Zhou J, Liew A W C, et al. CRF learning with CNN features for hyperspectral image segmentation[C]//2016 IEEE International Geoscience and Remote Sensing Symposium. Beijing, China.: 6890-6893
- [25] Zuo Zongcheng, Zhang Wen, Zhang Dongying. A Remote Sensing Image Semantic Segmentation Method by Combining Deformable Convolution with Conditional Random Fields[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(6): 718-726 (左宗成, 张文, 张东映. 融合可变形卷积与条件随机场的遥感影像语义分割方法[J]. *测绘学报*, 2019, 48(6): 718-726)
- [26] Teichmann M, Cipolla R. Convolutional CRFS for Semantic Segmentation[J]. *ArXiv Preprint ArXiv: 1805.04777*, 2018
- [27] Li Q Y, Shi Y L, Huang X, et al. Building Footprint Generation by Integrating Convolution Neural Network with Feature Pairwise Conditional Random Field (FPCRF)[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(11): 7502-7519
- [28] Li S Z. Markov random field modeling in image analysis[M]. 3rd ed. London: Springer, 2009
- [29] Adams A, Baek J, Davis M A. Fast High-Dimensional Filtering Using the Permutohedral Lattice[J]. *Computer Graphics Forum*, 2010, 29(2): 753-762

Road Extraction from Remote Sensing Image by Integrating DCNN with Short Range Conditional Random Field(SRCRF)

HE Xiaohui^{1,3} CHEN Mingyang² LI Panle² TIAN Zhihui^{1,3} ZHOU Guangsheng^{1,3}

1. School of Earth Science and Technology, Zhengzhou University, Zhengzhou 450052, China

2. School of Information Engineering, Zhengzhou University, Zhengzhou 450052, China

3. Joint Laboratory of Eco-Meteorology, Chinese Academy of Meteorological Sciences, Zhengzhou University, Zhengzhou 450052

Abstract: Objective: Deep Convolutional Neural Network (DCNN) is widely used in automatic road extraction

from high-resolution Remote Sensing images (HRSIs). However, the existing methods are difficult to model the context relationship between pixels in the predicted results. To solve this problem, some studies have used Fully Connected Crf (FullCrf) to perform secondary optimization of semantic segmentation results combined with context information, but the discontinuity problem of road structure cannot be effectively improved. In order to improve the integrity of road structure, this study proposes a Short Range Conditional Random Filed (SRCRF) model combined with DCNN. Methods: SRCRF mainly includes unary potential function based on road pre-segmentation, binary potential function based on spectral spatial features and k-neighborhood mean field inference algorithm. Firstly, the priori knowledge of road pre-segmentation results is obtained by using the powerful feature extraction capability of DCNN as the unary potential function of SRCRF, Secondly, the dependence of the binary potential function defined by the linear combination of Gaussian kernel functions on the surrounding nodes is modeled. The binary potential function enables the classification results to have local consistency that is, adjacent pixels with similar spectral features have the same label. Finally, K-neighborhood mean field inference algorithm based on mean field approximation inference algorithm optimizes the inference range to make full use of the spatial context information and spectral feature context information of the road, and then calculates the optimal label corresponding to each pixel based on the space and spectral feature to optimize the road accurately. The convolution method is adopted to control the inference range of SRCRF within the radius of K in order to improve the proportion of feature vectors (road-road). Results: The experimental results show that SRCRF alleviates the transition smoothness of FullCrf, and alleviates the structural discontinuity and incompleteness in the road acquisition results of high resolution remote sensing images. In Zimbabwe-Roads dataset and Cheng-Roads dataset, F1 values of SRCRF increased by about 4.01% and 3.73% respectively compared with DCNN, and about 3.25% and 2.28% respectively compared with FullCrf. Conclusion: Compared with traditional deep learning methods for road extraction from high resolution remote sensing images, this paper proposes a new road extraction scheme for remote sensing images, SRCRF. This scheme combines the advantages of DCNN and optimizes the fully connected structure of traditional conditional random fields into k-neighborhood structure, which reduces the inference scope and improves the proportion of feature vectors (road to road). Compared with FullCrf, SRCRF can make better use of image color features and spatial features to accurately optimize the road extraction results of deep learning output. According to the results of SRCRF on Cheng-Roads and Zimbabwe-Roads data sets, the performance of this method is improved compared with DCNN and FullCrf, and the time is shortened by one order of magnitude compared with FullCrf. In future work, we will further investigate the potential for learning Gaussian features and investigate more complex CRF architectures to better capture global context information. Finally, we are particularly interested in exploring the application potential of SRCRF in other fields, such as building extraction, vehicle extraction, lake extraction, etc.

Key words: Remote sensing image; Road extraction; Deep learning; Conditional random field; Semantic segmentation;

First author: HE Xiaohui, PhD, specializes in artificial intelligence, remote sensing image processing and data mining. E-mail: hexh@zzu.edu.cn

Foundation support: The Second Comprehensive Investigation and Research Project on Qinghai-Tibet Plateau.(2019QZKK0106)

网络首发:

标题: 结合DCNN与短距条件随机场的遥感影像道路提取

作者: 赫晓慧, 陈明扬, 李盼乐, 田智慧, 周广胜

DOI: 10.13203/j.whugis20210464

收稿日期: 2021-06-24

引用格式:

赫晓慧, 陈明扬, 李盼乐, 等. 结合DCNN与短距条件随机场的遥感影像道路提取[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210464 (HE Xiaohui, CHEN Mingyang, LI Panle, et al. Road Extraction From Remote Sensing Image by Integrating DCNN With Short Range Conditional Random Field(SRCRF)[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210464)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

全连接条件随机场高分辨率遥感影像面状地物交互提取

张春森, 胡艳, 于振, 崔卫红, 吴蓉蓉

武汉大学学报·信息科学版, 2020, 45(10): 1594-1601

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20190137>