

# CNN-ISS 遥感影像分类的瓦片边缘效应及消除方案

段增强, 刘杰东, 鹿 鸣, 孔祥斌, 杨 娜

(中国农业大学土地科学与技术学院 自然资源部农用地质量与监测重点实验室, 北京 100193)

**摘 要:** 应用卷积神经网络语义分割模型 (Image Semantic Segmentation based on Convolutional Neural Network, CNN-ISS) 进行遥感影像分类时, 需将大幅影像分解为特定大小瓦片影像, 并将其作为 CNN-ISS 处理对象, 这一过程破坏了位于瓦片边缘处地物的完整几何及纹理特征, 从而影响瓦片边缘处地物的识别效果, 即瓦片边缘效应。该研究以 DeepLab V3 为 CNN-ISS 核心模型, 对唐山农村地物进行语义分割, 定量分析了分类结果的瓦片边缘效应, 并提出了 5 个消除此效应的后处理方案。结果表明: 像素分类精度与像素到瓦片边缘距离正相关, 瓦片边缘处错误率最高达 6.93%, 中央处错误率最低为 3.52%, 存在瓦片边缘效应; 采用该研究提出的瓦片边缘效应消除方案后, 整幅影像的总精度 (Pixel Accuracy, PA)、均交并比 (Mean Intersection over Union, mIoU) 和 Kappa 系数均有提升, 最高分别提升 0.40、1.97 个百分点和 0.012 2。在不改变 CNN-ISS 核心模型条件下, 通过该研究的瓦片边缘效应消除后处理方案, 可有效提升遥感影像分类精度, 尤其针对复杂异构体和线状地物精度提升效果更好。

**关键词:** 遥感; 卷积神经网络; 语义分割; 影像分类; 瓦片边缘效应

doi: 10.11975/j.issn.1002-6819.2021.2.024

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2021)-2-0209-09

段增强, 刘杰东, 鹿鸣, 等. CNN-ISS 遥感影像分类的瓦片边缘效应及消除方案[J]. 农业工程学报, 2021, 37(2): 209-217. doi: 10.11975/j.issn.1002-6819.2021.2.024 http://www.tcsae.org

Duan Zengqiang, Liu Jiedong, Lu Ming, et al. Tile edge effect and elimination scheme of image classification using CNN-ISS remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(2): 209-217. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.2.024 http://www.tcsae.org

## 0 引 言

图像语义分割 (Image Semantic Segmentation, ISS) 是指, 为图像中的每一个像素分配一个预先定义好的表示其语义类别的标签<sup>[1]</sup>。基于卷积神经网络的图像语义分割方法 (Image Semantic Segmentation based on Convolutional Neural Network, CNN-ISS) 是一门涉及计算机视觉、模式识别与人工智能等研究领域的交叉学科, 是当前数字图像处理与机器视觉的研究热点内容之一<sup>[2-5]</sup>。作为典型 CNN-ISS 模型, DeepLab V3 模型在 Pascal VOC 数据集上的均交并比 (Mean Intersection over Union, mIoU) 值达到 86.9%, DeepLab V3+模型的分割精度达到了 89.0%, 在 Cityspace 公开数据集上也达到 82.1% 的分类精度<sup>[6-7]</sup>。CNN-ISS 在遥感影像分类方面也具有较为广泛的应用, 相较于传统遥感影像分类方法<sup>[8-9]</sup>, CNN-ISS 具有较强的迁移学习能力和泛化能力, 能够有效提取地物纹理、几何等深层次特征, 在高分辨率影像解译、复杂地物识别和作物分布制图等方面具备一定优势<sup>[10-14]</sup>。

出于效率或者硬件性能限制的考虑, 通常需要对大

幅的遥感影像进行瓦片切图<sup>[15]</sup>, 将生成的影像瓦片作为语义分割数据处理对象。人为的遥感影像瓦片化过程会产生瓦片边缘地物碎片, 进而影响遥感影像边缘附近像素的分类精度, 从而产生瓦片边缘附近像素分类精度低于瓦片中央区域像素分类精度的现象, 即瓦片边缘效应<sup>[16]</sup>。目前学者针对 CNN-ISS 遥感影像分类的研究主要有 3 个方面, 一是仅在瓦片尺度进行模型验证和精度评价<sup>[17-18]</sup>; 二是侧重于优化地物边缘分类的研究, 并未对瓦片边缘效应做专门分析及特别考虑<sup>[19-20]</sup>; 三是通过“重叠预测”和概率图累加算法来降低或消除瓦片边缘效应<sup>[21]</sup>。目前鲜见对 CNN-ISS 模型分类结果的瓦片边缘效应进行量化分析的研究文献。

本研究系统分析了研究区在高分辨率遥感影像 CNN-ISS (以 DeepLab V3 为核心模型) 分类中的瓦片边缘效应, 构建了一个消除 CNN-ISS 瓦片边缘效应的数据后处理框架, 以实现大幅遥感影像分类解译和瓦片边缘效应消除。

## 1 研究区概况

研究区位于河北省唐山市曹妃甸区北部以及滦南县南部地区 (118°18'E~118°48'E, 39°30'N~40°12'N)。具有典型的乡村土地覆被特点, 土地覆被类型主要包括水田、水浇地、设施农用地、农村居民点、河流、坑塘、交通用地等 (图 1)。

收稿日期: 2020-09-27 修订日期: 2021-01-03

基金项目: 自然资源部国土卫星遥感应用中心, 自然资源督察遥感监测指标与分析方法研究 (2020072; 11910661); 国家社会科学基金重大项目“休养生息制度背景下的耕地保护转型研究 (19ZDA096)

作者简介: 段增强, 副教授, 研究方向为土地信息技术。

Email: duanzq@cau.edu.cn

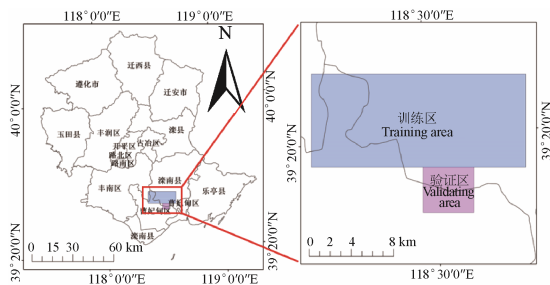


图 1 研究区地理位置  
Fig.1 Geographical location of the study area

2 数据与方法

2.1 数据源及其处理

2.1.1 训练区数据集构建

本研究训练区面积为 200 km<sup>2</sup>(图 1),研究采用 2019 年 5 月 4 日高分二号影像,经预处理后,最终得到空间分辨率为 1 m 的真彩色融合影像。通过人工目视解译土地覆被类型,分类标准如表 1 所示,进而构建 CNN-ISS 训练样本数据集(图 2)。

| 表 1 土地覆被类型及目视解译划分标准              |                                |
|----------------------------------|--------------------------------|
| 土地覆被类型 Land cover type           | 解译标准 Interpretation standards  |
| 农村居民点<br>Rural residential areas | 包括村庄内部宅基地、绿化林、空地、道路、环村庄道路及硬化地面 |
| 道路 Roads                         | 宽度大于 15 像素的公路                  |
| 河流沟渠 Rivers and ditches          | 宽度大于 30 像素的河流水面及沟渠             |
| 坑塘 Pit ponds                     | 坑塘水面以及塘基                       |
| 其他农用地<br>Other agricultural land | 耕地、设施农用地(主要为大棚)、林地             |

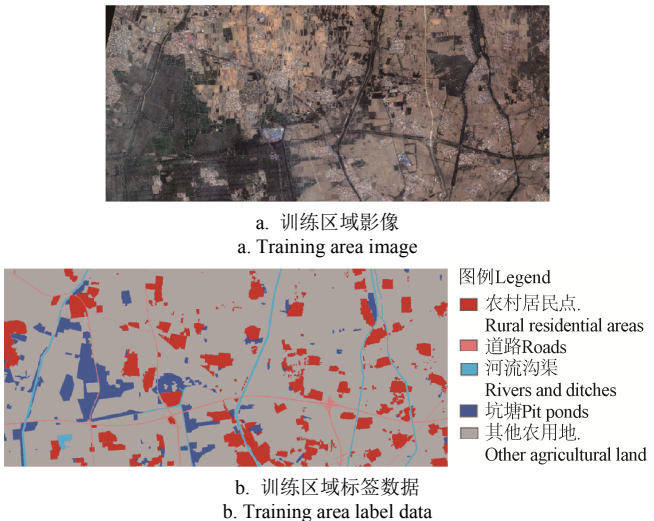


图 2 训练区域影像及对应的标签数据  
Fig.2 Training area image and corresponding label data

CNN-ISS 模型神经网络层数多、参数规模大,模型训练计算量大,出于模型学习精度和学习效率的考虑,通常限定特定尺寸图像作为训练和预测对象,常用的瓦片大小有 256×256、512×512 像素等。如果对大幅遥感影像进行语义分割,则需要首先将大幅影像进行瓦片化处理,以特定大小的瓦片作为基本处理单元。

本研究以 DeepLab V3 模型作为核心 CNN-ISS 模型,以 512×512 像素瓦片作为训练和预测处理单元。通过对

训练区顺序裁剪生成 2000 张训练瓦片,并在此基础上采用随机翻转、增加点噪声、高斯滤波和双边滤波等数据增强处理方式<sup>[22-23]</sup>,共计生成 15 000 幅瓦片。由于道路和河流沟渠像素占试验区影像总像素比例较低,采用过采样方案<sup>[24]</sup>增加了 3000 幅包括高速公路和河流瓦片。最终的训练集包括 512×512 像素瓦片共计 18 000 幅。选取其中的 75%作为训练数据集,25%作为验证数据集。综合考虑效率及最终精度,本试验以 ResNet V2\_50 预训练模型<sup>[25]</sup>为基础进行迁移训练。

2.1.2 预测区数据集构建

选取同幅影像内与训练区毗邻的 14 km<sup>2</sup> 区域作为预测区(图 3),采用与训练区相同的分类标准,通过人工目视解译对预测区土地覆被类型进行划分。

预测区影像、土地覆被类型划分及瓦片化结果如图 3 所示。模型精度评估、影像瓦片边缘效应分析与消除等均基于预测区数据集完成。

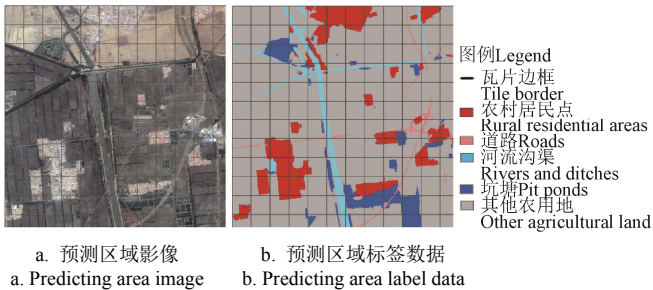


图 3 预测区域影像以及对应的标签瓦片数据  
Fig.3 Predicting area image and corresponding tile label data

2.2 研究方法

2.2.1 核心网络

语义分割(ISS)是在像素级别进行分类。CNN-ISS 将输入的遥感影像通过 CNN 提取影像特征,得到像素类别得分图后,再通过 softmax 函数以及 argmax 函数得到输入影像的像素类别<sup>[26]</sup>。

典型的语义分割模型数据处理过程<sup>[27]</sup>如图 4 中 CNN-ISS 标注部分所示,具体可以概况为以下 3 个过程:

1) 将瓦片输入至 CNN-ISS 模型中,进行影像特征提取。其公式为  $V=F(T,\theta)$ , 其中,  $V$  为经过系列特征提取后形成的类别得分矩阵,  $T$  为瓦片处理单元矩阵,  $\theta$  为 CNN-ISS 模型中各层网络的参数。实现过程依具体 CNN-ISS 模型而定,以 DeepLab V3 为例,这个过程包括了卷积、池化、空洞空间池化金字塔及上采样等数据处理过程。

2) 将类别得分图中各像素的分值分别经 softmax 函数处理后,统一归一化至 0~1 范围内,生成的图像为类别概率图。这一过程的公式为  $P=\text{softmax}(V)$ ,  $P$  为类别概率矩阵。

3) 选取某一像素点对应的概率矩阵中最大值的索引作为该像素的最终类别,这一过程输出的结果即为遥感影像分类图。这一过程的公式为  $C=\text{argmax}(P)$ ,  $C$  为分类结果矩阵。

本研究以 DeepLab V3 作为核心模型,该模型以 ResNet 50 作为特征提取网络,加入了空洞空间池化金字

塔 (Atrous Spatial Pyramid Pooling, ASPP) 模块来提取全局、多尺度信息。ASPP 采用了 4 个并行的卷积, 即 1 个  $1 \times 1$  的卷积和 3 个扩张率分别为 (6、12、18) 的  $3 \times 3$

扩张卷积, 同时加入了批归一化 (Batch Normalization, BN) 层对影像特征进行处理, 最后通过 16 倍的上采样恢复到原图大小<sup>[28]</sup>。

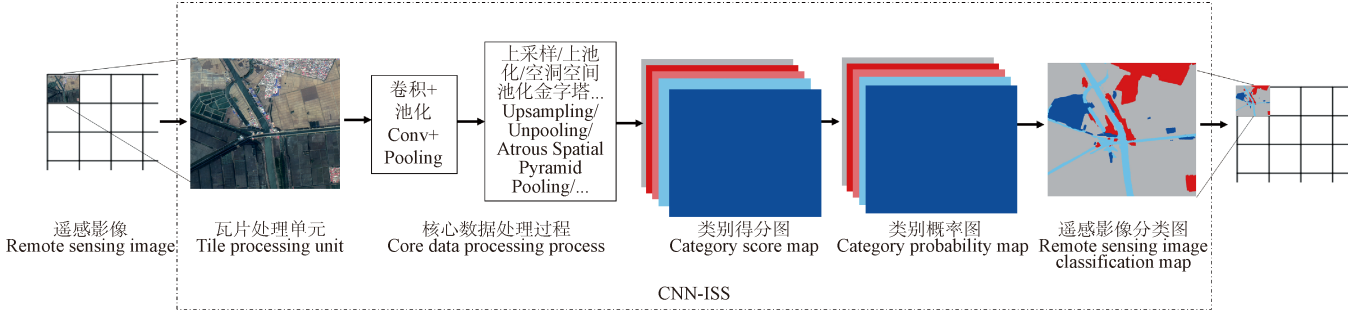


图 4 CNN-ISS 模型示意图

Fig.4 Diagram of Image Semantic Segmentation based on Convolutional Neural Network (CNN-ISS) model

### 2.2.2 瓦片边缘效应消除方案

遥感影像瓦片化会在瓦片边缘产生土地覆被类型碎片, 从而影响分类精度, 存在瓦片边缘效应。

本研究中, 以整幅影像上做移窗切图后生成的  $512 \times 512$  像素瓦片影像作为处理单元, 纵向和横向均以 512 像素步长, 依次从左到右, 从上到下进行移窗和瓦片影像语义分割。为消除瓦片边缘效应, 设定不同的移窗起点位置  $(i, k)$ ,  $i$ 、 $k$  为该移窗偏置起点的像素行列位置, 从而保证整幅遥感影像上任意像素点一定处于某次偏置设置下生成的瓦片中央区域。假定共进行  $m$  次移窗, 则可为每个输入影像像素生成至多  $m$  批次的类别得分值, 假定有  $n$  个分类类别, 则每个像素至多生成  $m \times n$  个类别得分值, 以该像素的  $m \times n$  个类别得分值为基础, 设计不同的瓦片效应消除技术方案, 从而提高语义分割精度。

设输入遥感影像为  $H \times W \times 3$  的矩阵  $X$ , 其中,  $H$  为影像高度,  $W$  为影像宽度,  $H$  和  $W$  单位都为像素, 像素  $x_{(a,b)} = [x_R, x_G, x_B]$ ,  $x_R, x_G, x_B$  代表输入影像的彩色通道,  $(a, b)$  代表某一像素在输入遥感影像上的坐标。

$$X = \begin{pmatrix} x_{(0,0)} & \dots & x_{(0,W)} \\ \dots & \dots & \dots \\ x_{(H,0)} & \dots & x_{(H,W)} \end{pmatrix} \quad (1)$$

对  $X$  设定一个移窗切图的移窗起点为  $(i, k)$ , 进行移窗切图, 将生成的瓦片数据集经核心 CNN-ISS 模型处理得到瓦片像素类别得分结果, 拼接后得到全图得分结果, 记为类别得分矩阵  $V_{(i,k)}$ , 其中  $(i, k)$  为偏置起点的像素位置。对偏置起点为  $(0, 0)$  的情况, 得到类别得分矩阵  $V_{(0,0)}$ 。

$$V_{(0,0)} = \begin{pmatrix} v_{(0,0)} & \dots & v_{(0,W)} \\ \dots & \dots & \dots \\ v_{(H,0)} & \dots & v_{(H,W)} \end{pmatrix} \quad (2)$$

其中  $v_{(a,b)}$  为像素  $x_{(a,b)}$  的各类别得分值所组成的一维向量组, 即  $v_{(a,b)} = [v_{1(a,b)}, \dots, v_{c(a,b)}, \dots, v_{n(a,b)}]$ ,  $v_{c(a,b)}$  代表像素  $x_{(a,b)}$  预测分类  $c$  的类别得分值,  $c=1, 2, \dots, n$ ,  $n$  为语义分割的分类数量。

当偏置起点为  $(i, k)$  时, 输入影像的前  $i$  行和前  $k$  列像素不参与瓦片生成和后续的 CNN-ISS 核心模型数据处理, 因此无法得到其所对应的类别得分, 则为其进行非

数填充处理, 在下边的矩阵表达中用  $NaN$  表示, 得到偏置起点为  $(i, k)$  的类别得分矩阵  $V_{(i,k)}$

$$V_{(i,k)} = \begin{pmatrix} NaN & \dots & NaN & \dots & NaN \\ \dots & \dots & \dots & \dots & \dots \\ NaN & \dots & v_{(i,k)} & \dots & v_{(i,W)} \\ \dots & \dots & \dots & \dots & \dots \\ NaN & \dots & v_{(H,k)} & \dots & v_{(H,W)} \end{pmatrix} \quad (3)$$

假定第  $s$  次偏置中, 像素的  $c$  类别得分值记为  $v_{c(a,b)}^s$ , 为了消除基于瓦片数据处理而产生的瓦片边缘效应, 本研究进行多次、不同的偏置处理, 并对多次偏置生成的多个类别得分矩阵进行合成。  $v_{c(a,b)}^s$  和  $p_{c(a,b)}^s$  表示编号  $s$  的偏置方案中, 像素  $x_{(a,b)}$  对应类别  $c$  的类别得分值和概率值,  $v_{c(a,b)}$  和  $p_{c(a,b)}$  则表示合成后的  $c$  类别得分值和概率值;  $c_{(a,b)}$  表示像素  $x_{(a,b)}$  的分类结果。其中,  $c \in [1, n]$ ,  $s \in [1, m]$ 。具体的瓦片边缘效应消除方案处理逻辑如表 2 所示, 其操作对象均为像素  $x_{(a,b)}$ , 故将下标  $(a, b)$  进行省略。

表 2 瓦片边缘效应消除方案

| 方案 Solution        | 处理逻辑 Processing logic                                                                                                                                   |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| 方案 1<br>Solution 1 | 1) $s = \text{argmax}(d^s \mid s \in [1, m])$<br>2) $p_c = \text{softmax}(v_c^s), c=1, 2, \dots, n$<br>3) $c = \text{argmax}(p_c \mid c \in [1, n])$    |
| 方案 2<br>Solution 2 | 1) $v_c = \max(v_c^s), s=1, 2, \dots, m$<br>2) $p_c = \text{softmax}(v_c), c=1, 2, \dots, n$<br>3) $c = \text{argmax}(p_c \mid c \in [1, n])$           |
| 方案 3<br>Solution 3 | 1) $v_c = \text{arv}(v_c^s), s=1, 2, \dots, n$<br>2) $p_c = \text{softmax}(v_c), c=1, 2, \dots, n$<br>3) $c = \text{argmax}(p_c \mid c \in [1, n])$     |
| 方案 4<br>Solution 4 | 1) $p_c^s = \text{softmax}(v_c^s), s=1, 2, \dots, m$<br>2) $p_c = \max(p_c^s), c=1, 2, \dots, n$<br>3) $c = \text{argmax}(p_c \mid c \in [1, m])$       |
| 方案 5<br>Solution 5 | 1) $p_c^s = \text{softmax}(v_c^s), c=1, 2, \dots, n$<br>2) $p_c = \text{arv}(p_c^s), s=1, 2, \dots, m$<br>3) $c = \text{argmax}(p_c \mid c \in [1, n])$ |

注:  $d^s$  为第  $s$  次偏置时像素  $x_{(a,b)}$  到瓦片边缘的距离,  $s \in [1, m]$ 。

Note:  $d^s$  is the distance from the pixel  $x_{(a,b)}$  to the edge of the tile at the  $s$ -th bias,  $s \in [1, m]$ .

方案 1: 假设影像中某一点在第  $s$  次偏置时, 到瓦片

边缘的距离最大,则选取此次偏置中该点对应的打分值作为该点的最终打分值  $v_c^s$ ,后经 softmax 函数处理得到整幅影像的类别概率值  $P_c$ ,进而得到最终分类结果。

方案 2:对每个类别  $c$ ,在  $m$  次偏置中形成的  $m$  个像素类别打分值  $v_c^s$  中选最大值,作为类别  $c$  的类别判断打分值  $v_c$ ,进而根据  $v_c$  计算类别概率  $p_c$  和分类。

方案 3:对  $m$  次偏置中形成的  $m$  个像素类别打分值  $v_c^s$  求平均值,作为类别  $c$  的类别判断打分值  $v_c$ ,通过 softmax 函数得到类别概率  $p_c$ ,然后选取  $p_c$  对应的类别作为最终分类结果。

方案 4:将  $m$  次移窗得到的  $m$  组打分值分别进行 softmax 处理,得到  $m$  组像素类别概率值  $p_c^s$ ,选择其中最大的类别概率值作为类别  $c$  的类别判断概率值  $p_c$ ,并将  $p_c$  对应的类别作为像素的最终类别。

方案 5:对于 softmax 得到的  $m$  组像素类别概率值  $p_c^s$  取平均值,将该值作为类别  $c$  的类别判断概率值  $p_c$ ,进而得到像素的最终分类结果。方案 5 将多次移窗的类别判别概率的平均值作为最终的类别判别概率值。陈守玉<sup>[21]</sup>将多次移窗类别判别概率值的加和值作为最终的类别判别概率值。因此本研究中方案 5 和陈守玉的处理方式在本质上是一致的。

需要指出的是,本研究通过消除瓦片边缘效应来提升语义分割精度,是对 CNN-ISS 瓦片化处理结果的一种后处理方案,可采用不同的核心模型,且可以直接继承所采用不同核心模型的预训练结果或迁移学习结果。

### 2.2.3 精度评价指标

本研究采用总精度 (Pixel Accuracy, PA)<sup>[29]</sup>、Kappa 系数<sup>[30]</sup>、交并比 (Intersection over Union, IoU)<sup>[31]</sup>、均交并比 (Mean Intersection over Union, mIoU)<sup>[32]</sup>作为精度评价指标。PA、Kappa 系数、IoU 和 mIoU 的数值越大,说明 CNN-ISS 模型分类精度越高。

瓦片边缘效应具体体现为距离瓦片边缘愈近的区域,其错分像素占比愈高。本研究通过构建瓦片边缘距离误判率 ERD (Error Rate with a Distance to tile edges, ERD) 和整体误判率 ERW (Error Rate of the Whole image, ERW) 2 个指标,来对 CNN-ISS 遥感影像分类解译的瓦片边缘效应进行定量分析,其计算式如下

$$ERD(d) = \frac{N_d^F}{N_d} \quad (4)$$

$$ERW = 1 - PA = \frac{N^F}{N} \quad (5)$$

式中  $d$  记为到瓦片边缘的像素距离,单位为像素,  $N$  为栅格总数,  $N_d$  瓦片边缘距离为  $d$  的栅格总数,  $N_d^F$  表示到瓦片边缘距离为  $d$  的错分栅格总数,  $N^F$  表示错分栅格总数。

## 3 结果与分析

### 3.1 DeepLab V3 分类结果分析

本研究以 DeepLab V3 作为 CNN-ISS 核心模型。

偏置起点设为(0,0),在不进行多次偏置前提下,得到的影像分类结果等同于将 DeepLab V3 原始模型分类瓦片直接拼接形成的分类图。本研究基于此结果对 DeepLab V3 分类精度进行评估,同时将其作为边缘效应消除方案的对照数据。由于各瓦片边缘效应消除方案均设置有多个瓦片切割起点偏置,因此有效预测范围与对照组不同,以下所有分析均选取公共有效区域进行。

本试验操作系统为 Windows10,编程语言为 Python 3.5,使用开源深度学习框架 TensorFlow 1.9.0,计算机 CPU 为 i7-7700HQ,配 16 GB RAM 和一块 GTX 1070 (8 GB)。DeepLab V3 模型训练参数 batch\_size 为 4,初始学习率为 0.000 1,每 8 000 次迭代学习率递减 50%,L2 正则化系数为 0.000 1。

在最终的分类模型中,训练集 loss 值为 0.069 8,验证集 loss 值为 0.001 1,在预测区总分类精度为 94.99%,Kappa 系数为 0.868 8,mIoU 值为 76.24%,各覆盖类型 IoU 值见表 3 中的对照组数据。与相关研究结果对比<sup>[13]</sup>,本研究的 PA、Kappa 系数和 mIoU 值均达到较高数值,说明本研究的 CNN-ISS 模型具有较高分类精度,但由于训练数据及数据量较小,且道路和沟渠占比较少,导致其 IoU 相对较低。

表 3 对照组精度汇总表  
Table 3 Accuracy summary table of control group

| 指标 Index                                      | 数值 Value    |
|-----------------------------------------------|-------------|
| 交并比<br>Intersection over Union (IoU)/%        | 其他农用地 94.35 |
|                                               | 农村居民点 88.39 |
|                                               | 道路 53.33    |
|                                               | 河流沟渠 71.66  |
|                                               | 坑塘 73.47    |
| 总精度 Pixel Accuracy (PA) /%                    | 94.99       |
| Kappa 系数 Kappa coefficient                    | 0.868 8     |
| 均交并比<br>Mean Intersection over Union (mIoU)/% | 76.24       |

CNN-ISS 分类精度与标签精度及训练样本数据量有关,相关研究已经验证了 DeepLab V3 模型在遥感影像分类中的性能<sup>[2,33-36]</sup>,而本研究重点在于瓦片边缘效应分析及其消除方案,因此不再对 DeepLab V3 在本试验中各地类精度进行更详细分析。

### 3.2 瓦片边缘效应分析

当遥感影像中的地物位于输入模型瓦片的边缘处时,会产生地物碎片或不完整地物,从而丢失部分识别信息,影响到分类精度。

本研究 CNN-ISS 影像瓦片大小为 512×512 像素,预测区总面积为 14 km<sup>2</sup>。将每个瓦片按照 3×3 平均划分为 9 个区域,并将外围的 8 个子区作为边缘区域,中央子区作为中央区域,分别以边缘区域、中央区域和整个瓦片区域作为统计范围。14 km<sup>2</sup> 预测区内对照组所有瓦片边缘处与中心处的精度评估结果如表 4 所示,整个瓦片区域的精度评价结果如表 3 所示。

表 4 对照组瓦片边缘与中心处精度对比  
Table 4 Accuracy comparison between tip edge and center of control group

| 统计范围<br>Statistical scope | IoU/%                            |                                  |             |                            |                 | PA/%  | ERW/% | Kappa 系数<br>Kappa coefficient | mIoU/% |
|---------------------------|----------------------------------|----------------------------------|-------------|----------------------------|-----------------|-------|-------|-------------------------------|--------|
|                           | 其他农用地<br>Other agricultural land | 农村居民点<br>Rural residential areas | 道路<br>Roads | 河流沟渠<br>Rivers and ditches | 坑塘<br>Pit ponds |       |       |                               |        |
| 边缘区域 Edge area            | 94.27                            | 88.22                            | 53.26       | 72.02                      | 72.80           | 94.91 | 5.09  | 0.867 3                       | 76.11  |
| 中央区域 Central area         | 94.94                            | 89.65                            | 53.94       | 68.25                      | 79.21           | 95.63 | 4.37  | 0.880 6                       | 77.20  |

注：ERW 为整体误判率。  
Note: ERW represents error rate of the whole image.

PA、Kappa 和 mIoU 评估结果整体表现为中央区域>整个瓦片区域>边缘区域；根据各地类中央区域 IoU 与边缘 IoU 的差值从高到低排序依次为，坑塘（6.41 个百分点）>农村居民点（1.42 个百分点）>道路（0.68 个百分点）>其他农用地（0.67 个百分点）>河流沟渠（-3.77 个百分点），其中坑塘、农村居民点中央区域 IoU 与边缘区域 IoU 相比均高出 1%以上，说明这 2 个土地覆盖类型的边缘效应更加突出。对照遥感影像和预测区标签可以发现，坑塘和农村居民点均由多种颜色和纹理的小斑块组成，属于典型的“异构体”复杂对象，因此瓦片切割产生的碎片更易损失对象的整体特征信息，因而其瓦片边缘效应更加突出。当道路、河流沟渠等线状地物靠近瓦片边缘或以较小夹角通过瓦片边缘时，也较易产生琐碎或狭长碎片，从而也具有瓦片边缘效应。“其他农用地”是整个景观中“基质”成分，瓦片切割造成的独立和零散分布的碎片较少，瓦片边缘效应相对较弱，表 3 数据也表明其他农用地具有较弱的瓦片边缘效应。

道路和河流沟渠在训练区样本数量都比较低，包含道路和河流沟渠的瓦片数量少，包含道路和河流沟渠的瓦片中其像素占比也很小，造成道路和河流沟渠分类精度相对较低。在预测区也存在类似情况，使得道路和河流沟渠精度统计学意义降低。

3.3 瓦片边缘效应消除

为消除边缘效应，本研究设置的横、纵向移窗偏置步长均为 1/3 瓦片宽度（高度），共计 9 次偏置，即在本研究中， $s=9$ ，因此除整幅遥感影像边缘处外，可以确保任意像素一定处于某次偏置设置下生成的移窗瓦片的中央区域。

运用 2.2.2 中各方案对预测区影像进行重新分类，测试了不同瓦片边缘效应消除方案的效果。如图 5 所示，瓦片边框处的河流沟渠、农村居民点和坑塘，瓦片内地物碎片导致分类结果出现明显的局部错分或漏分。较对照组，本研究所提 5 种瓦片边缘消除方案对瓦片边缘处的地物分类效果均有不同程度提升。

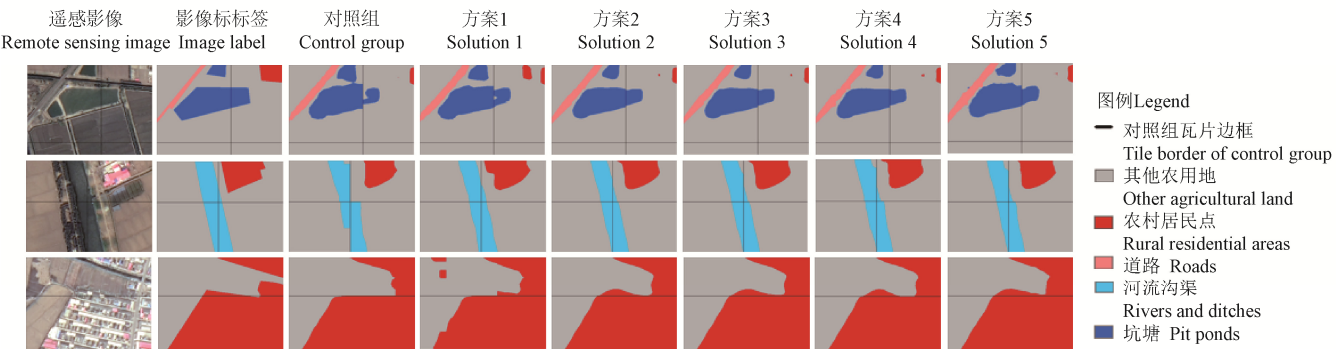


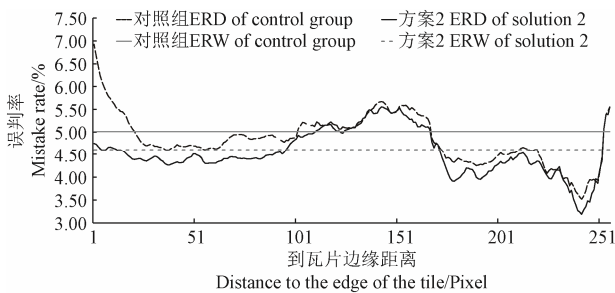
图 5 各方案局部示意图  
Fig.5 Partial schematic diagram of each scheme

各方案精度汇总如表 5 所示。由表 5 可知，相较于对照组，方案 1~5 的各项评价指标均有不同程度提升。不同方案的 Kappa 系数从高到低依次为：方案 2（0.881 0）>方案 5（0.878 9）>方案 3（0.878 8）>方案 4（0.877 7）>方案 1（0.875 9），均大于对照组（0.868 8）。方案 2 的分类效果最好，因此本文将以方案 2 为例，分析不同土地覆被类别精度可知，相对于对照组，各类型 IoU 提升幅度次序从高到低为：道路（4.13 个百分点）>坑塘（2.97 个百分点）>河流沟渠（1.61 个百分点）>农村居民点（0.65 个百分点）>其他农用地（0.46 个百分点），总体规律表现为线状地物和复杂异构体（坑塘、农村居民点）的提升幅度高于基底地类（其他农用地）。方案 2 的 mIoU 最高达到 78.20%，较对照组的 76.24%提升了 1.97 百分点。

表 5 各方案精度汇总表  
Table 5 Accuracy summary table of each solution

| 指标 Index                      | 方案 1<br>Solution 1 | 方案 2<br>Solution 2 | 方案 3<br>Solution 3 | 方案 4<br>Solution 4 | 方案 5<br>Solution 5 |
|-------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 其他农用地                         | 94.64              | 94.81              | 94.73              | 94.70              | 94.74              |
| 农村居民点                         | 88.92              | 89.04              | 89.03              | 88.84              | 89.08              |
| 道路                            | 54.75              | 57.46              | 54.58              | 54.54              | 54.69              |
| 河流沟渠                          | 72.90              | 73.27              | 74.79              | 74.09              | 74.61              |
| 坑塘                            | 75.17              | 76.44              | 75.71              | 75.56              | 75.72              |
| PA/%                          | 95.25              | 95.39              | 95.38              | 95.32              | 95.38              |
| ERW/%                         | 4.75               | 4.61               | 4.62               | 4.68               | 4.62               |
| Kappa 系数<br>Kappa coefficient | 0.875 9            | 0.881 0            | 0.878 8            | 0.877 7            | 0.878 9            |
| mIoU/%                        | 77.28              | 78.20              | 77.77              | 77.55              | 77.77              |

预测区内对照组和方案 2 的所有瓦片的瓦片边缘距离误判率 ERD 统计结果如图 6 所示。对照组瓦片边缘距离误判率 ERD 随着到瓦片边缘距离的增加呈现波动下降的趋势,在瓦片边缘处时,ERD 最高,达到了 6.93%,在到瓦片边缘距离为 242 像素时,RED 最低,为 3.52%,而整体误判率 ERW 则为 5.01%。ERD 与到瓦片边缘的距离负相关,具有瓦片边缘效应。对照组整体误判率 ERW 为 5.01%,方案 2 ERW 下降到 4.61%。与对照组相比,方案 2 ERD 在所有距离上均有所下降,其中瓦片边缘附近 ERD 值由对照组的 6.93% 下降至 4.75%,下降幅度最大,说明像素到边缘的距离对方案 2 ERD 的影响降低,瓦片边缘效应得到有效改善。



注: ERD 为瓦片边缘距离误判率。

Note: ERD represents error rate with a distance to tile edges.

图 6 错分栅格统计

Fig.6 Mismatched raster statistics

## 4 讨论与结论

### 4.1 讨论

本研究中 DeepLab V3 模型对高分辨率下具有复杂遥感影像特征的异构体(坑塘、农村居民点)识别能力较好,而对于线状地物(道路、河流沟渠)分辨能力较弱,产生这种现象的一个重要原因是训练区中的道路和河流沟渠样本占总像素比例低,即使进行过采样预处理,模型对于这种小样本地物的影像特征获取依然不足,这个问题在遥感影像 CNN-ISS 分类中普遍存在。因此可从以下几个方面进行优化:一是在数据预处理阶段,除了一般的图像处理数据增强手段,以及人为增加包含像素占比较少的小样本地物(如道路、河流沟渠等)瓦片数量的过采样方式,亦可通过“小样本地物提取+背景转换”的高级数据扩展方法,从而尽可能增加小样本地物、及其在各种环境场景中的图像特征信息<sup>[37]</sup>。二是构造适用于小样本数据的损失函数,通过增大损失函数中小样本地物的权重值来解决由于不平衡样本导致的模型精度问题。三是针对不同地物分类效果的差异特性,构建一种混合分类模型。例如,线状地物的像素在整幅影像中占比少且纹理简单,传统遥感影像分类方法对于此种类型地物识别效果较好,而对于复杂异构体而言, CNN-ISS 模型能够更为有效提取其深层次影像特征,在今后的研究中可考虑根据地物类型整合使用各自最优的分类模型。

本研究提出的通过多次移窗偏置消除瓦片边缘效应的技术方案,本质上是通过多视野的方法降低地物碎片

对图像分类精度的影响,而多视野信息是否综合以及如何综合消除瓦片边缘效应的效果各方案也有所不同。本研究提出的方案 1 直接将分类置信度最高视野中的像素分类作为该像素最终的分类结果,并未进行多视野信息综合;方案 2~5 不但进行多视野判断,还同时对每个视野(移窗)获取的类别打分矩阵或类别概率矩阵进行了信息综合。方案 2、3 分别在类别打分层次对多视野打分结果取最大值和取平均值,然后再以类别打分为基础计算类别概率值,进而对像素类别进行判断。方案 4、5 则在类别判断概率层次对多视野信息进行综合,其中方案 4 对多视野形成的类别判断概率取最大值,方案 5 对多视野类别判断概率取平均值;然后再在类别判断概率的基础进行类别判断。研究表明,上述多视野信息综合方案中,方案 2 效果分类精度最高。本研究提出的方案 1~5 的分类精度差异是否在不同场景保持一致有待进一步验证和分析。

对于具有代表性的 DeepLab 模型,从 DeepLab V3 到 DeepLab V3-JFT、DeepLab V3+ (Xception)、DeepLab V3+ (Xception-JFT)的迭代过程中,在 PASCAL VOC 2012 数据集的 mIoU 依次分别提升了 1.2%、0.9%和 1.2%<sup>[7]</sup>。Liu 等构建了一种全卷积神经网络——SRI-Net 模型,在 WHU Aerial Building Dataset 中得到的 IoU 值较 DeepLab V3+ 在该数据集中得到的 IoU 提升了 1.78%<sup>[38]</sup>。本研究提出的消除瓦片边缘效应的技术方案是对 CNN-ISS (DeepLab V3) 的后处理方案,以方案 2 为例,相较于对照组,其 mIoU 提升了 1.97 个百分点,特别是对于线状地物和异构体的提升效果更大,对道路和坑塘的 IoU 分别提升了 4.13 和 2.97 个百分点。因此研究者认为,本研究的提出的瓦片边缘效应消除方案,作为一种 CNN-ISS 模型的后处理方案,能够有效提升整幅遥感影像分类精度,尤其针对复杂异构体和线状地物精度提升幅度更大。但是上述结论还应该在更多应用场景中进一步验证。

本文所用到的高分二号影像,经过预处理后,其空间分辨率为 1m,在这种空间分辨率影像中,城市高大建筑物的阴影和侧墙面均表现明显,且对分类有较大影响。目前有研究提出专门的模型用于消除高大建筑物阴影对建筑物识别的影响<sup>[39]</sup>。城市建筑物,特别是具有复杂平面结构的建筑物,在瓦片化过程中的瓦片化效应应该更加突出,在未来可单独针对此应用场景展开瓦片边缘效应的相关研究。

### 4.2 结论

由瓦片边缘距离误判率的统计结果可知, CNN-ISS 分类精度与像素位置到瓦片边缘的距离呈正相关关系,瓦片中央区域精度高于瓦片边缘区域;各类型地物瓦片中央与边缘处的精度差异表明,不同类型地物的瓦片边缘效应程度不同,由复杂颜色、纹理图像特征构成的复杂异构体(农村居民点、坑塘)和线状地物瓦片边缘效应较其他地物的瓦片边缘效应更加突出。本文设计的 5 个消除瓦片边缘效应的技术方案,均可从不同程度消除瓦片边缘效应,其中取多次

偏置中各像素类别打分值最大值对应的类别作为其最终类别的方案对精度的提升最大;此外,瓦片边缘效应消除方案对不同类型地物的效果也有差异,瓦片边缘效应消除方案对复杂异构体(农村居民点,坑塘)和线状地物分类精度的提升更大。

#### [参 考 文 献]

- [1] Csurka G, Perronnin F. An efficient approach to semantic segmentation[J]. *Ernational Journal of Computer Vision*, 2011, 95(2): 198-212.
- [2] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. *软件学报*, 2019, 30(2): 440-468.  
Tian Xuan, Wang Liang, Ding Qi. A review of image semantic segmentation methods based on deep learning[J]. *Journal of Software*, 2019, 30(2): 440-468. (in Chinese with English abstract)
- [3] Zhang W, Li R, Deng H, et al. Deep convolucional neural networks for multi-modality isointense infant brain image segmentation[J]. *NeuroImage*, 2015, 108: 214-224.
- [4] Chen W, Zhang K, Liu Q, et al. Robust visual tracking via patch based kernel correlation filters with adaptive multiple feature ensemble[J]. *Neurocomputing*, 2016, 214: 607-617.
- [5] 陈鸿翔. 基于卷积神经网络的图像语义分割[D]. 杭州: 浙江大学, 2016.  
Chen Hongxiang. Image Semantic Segmentation based on Convolutional Neural Network[D]. Hangzhou: Zhejiang University, 2016. (in Chinese with English abstract)
- [6] Chen L, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[C]. *IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018.
- [7] Chen L, Zhu Y, Papandreou G, et al. Encoder-Decoder with atrous separable convolution for semantic image segmentation[C]// *European Conference on Computer Vision (ECCV)*. Cham, Springer, 2018: 833-851.
- [8] Zhang T, Su J, Liu C, et al. Potential bands of sentinel-2A satellite for classification problems in precision agriculture[J]. *International Journal of Automation and Computing*, 2019, 16(1): 16-26.
- [9] Sun Q, Zhang P, Wei H, et al. Improved mapping and understanding of desert vegetation-habitat complexes from intraannual series of spectral endmember space using cross-wavelet transform and logistic regression[J]. *Remote Sensing of Environment*, 2020, 236: 111516.
- [10] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(12): 2481-2495.
- [11] Gang F, Changjun L, Rong Z, et al. Classification for high resolution remote sensing imagery using a fully convolutional network[J]. *Remote Sensing*, 2017, 9(6): 498-515.
- [12] Zhou W, Newsam S, Li C, et al. Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval[J]. *Remote Sensing*, 2016, 9(5): 489-510.
- [13] Lv X, Ming D, Chen Y Y, et al. Very high-resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification[J]. *International Journal of Remote Sensing*, 2019, 40(2): 506-531.
- [14] 周壮, 李盛阳, 张康, 等. 基于 CNN 和农作物光谱纹理特征进行作物分布制图[J]. *遥感技术与应用*, 2019, 34(4): 694-670.  
Zhou Zhuang, Li Shengyang, Zhang Kang, et al. Crop distribution mapping based on CNN and crop spectral texture features[J]. *Remote Sensing Technology and Application*, 2019, 34(4): 694-670. (in Chinese with English abstract)
- [15] 杨建宇, 周振旭, 杜贞容, 等. 基于 SegNet 语义模型的高分辨率遥感影像农村建设用地提取[J]. *农业工程学报*, 2019, 35(5): 251-258.  
Yang Jianyu, Zhou Zhenxu, Du Zhenrong, et al. High-resolution remote sensing image rural construction land extraction based on SegNet semantic model[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2019, 35(5): 251-258. (in Chinese with English abstract)
- [16] Paton P W C. The effect of edge on avian nest success: How strong is the evidence?[J]. *Conservation Biology*, 1994, 8(1): 17-26.
- [17] 王俊强, 李建胜, 丁波, 等. 深度学习语义分割方法在遥感影像分割中的性能分析[J]. *计算机测量与控制*, 2019, 27(7): 231-235.  
Wang Junqiang, Li Jiansheng, Ding Bo, et al. Performance analysis of deep learning semantic segmentation methods in remote sensing image segmentation[J]. *Computer Measurement and Control*, 2019, 27(7): 231-235. (in Chinese with English abstract)
- [18] 李欣, 唐文莉, 杨博. 利用深度残差网络的高分遥感影像语义分割[J]. *应用科学学报*, 2019, 37(2): 282-290.  
Li Xin, Tang Wenli, Yang Bo. High-resolution remote sensing image semantic segmentation using deep residual network[J]. *Journal of Applied Sciences*, 2019, 37(2): 282-290. (in Chinese with English abstract)
- [19] 李雨慧. 基于深度学习的高分辨率遥感影像多尺度语义分割[D]. 上海: 上海交通大学, 2018.  
Li Yuhui. Multi-scale Semantic Segmentation of High-resolution Remote Sensing Images based on Deep Learning[D]. Shanghai: Shanghai Jiao Tong University, 2018. (in Chinese with English abstract)
- [20] 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.  
Zuo Zongcheng, Zhang Wen, Zhang Dongying. Remote sensing image semantic segmentation method combining

- deformable convolution and conditional random field[J]. Journal of Surveying and Mapping, 2019, 48(6): 718-726. (in Chinese with English abstract)
- [21] 陈守玉. 基于深度学习的遥感影像语义分割方法研究[D]. 哈尔滨: 东北林业大学, 2019.
- Chen Shouyu. Research on Remote Sensing Image Semantic Segmentation Method based on Deep Learning[D]. Harbin: Northeast Forestry University, 2019. (in Chinese with English abstract)
- [22] 李欣, 唐文莉, 杨博. 利用深度残差网络的高分遥感影像语义分割[J]. 应用科学学报, 2019, 37(2): 282-290.
- Li Xin, Tang Wenli, Yang Bo. High-Resolution remote sensing image semantic segmentation using deep residual network[J]. Journal of Applied Sciences, 2019, 37(2): 282-290. (in Chinese with English abstract)
- [23] Fortuna J, Martens H, Johansen T A. Multivariate image fusion: A pipeline for hyperspectral data enhancement[J]. Chemometrics and Intelligent Laboratory Systems, 2020, 205(15): 104097.
- [24] Iren Valova, Christopher Harris, Tony Mai, et al. Optimization of convolutional neural networks for imbalanced set classification[J]. Procedia Computer Science, 2020, 176: 660-669.
- [25] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition (CVPR). Las Vegas, Nevada: IEEE, 2016. 770-778.
- [26] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [27] Liangchih C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 40(4): 357-361.
- [28] 张鑫禄, 张崇涛, 戴晨光, 等. 基于 DeepLabv3 架构的高分辨率遥感图像分类[J]. 海洋测绘, 2019, 39(2): 40-44.
- Zhang Xinlu, Zhang Chongtao, Dai Chenguang, et al. High-resolution remote sensing image classification based on DeepLabv3 architecture[J]. Ocean Surveying and Mapping, 2019, 39(2): 40-44. (in Chinese with English abstract)
- [29] Foody G M. Status of land cover classification accuracy assessment[J]. Remote Sensing of Environment, 2002, 80(1): 185-201.
- [30] Persello C, Bruzzone L. A novel protocol for accuracy assessment in classification of very high-resolution images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 48(3): 1232-1244.
- [31] 盛恒, 黄铭, 杨晶晶. 基于 Faster R-CNN 和 IoU 优化的实验室人数统计与管理系统[J]. 计算机应用, 2019, 39(6): 1669-1674.
- Sheng Heng, Huang Ming, Yang Jingjing. Laboratory population statistics and management system optimized based on Faster R-CNN and IoU[J]. Computer Applications, 2019, 39(6): 1669-1674. (in Chinese with English abstract)
- [32] 冯鑫. 基于语义分割的行人属性识别技术的研究与实现[D]. 北京: 北京邮电大学, 2019.
- Feng Xin. Research and Implementation of Pedestrian Attribute Recognition Technology based on Semantic Segmentation[D]. Beijing: Beijing University of Posts and Telecommunications, 2019. (in Chinese with English abstract)
- [33] 陈前, 郑利娟, 李小娟, 等. 基于深度学习的高分遥感影像水体提取模型研究[J]. 地理与地理信息科学, 2019, 35(4): 43-49.
- Chen Qian, Zheng Lijuan, Li Xiaojuan, et al. Research on water extraction model of high-resolution remote sensing image based on deep learning[J]. Geography and Geo-Information Science, 2019, 35(4): 43-49. (in Chinese with English abstract)
- [34] 汪梓艺, 苏育挺, 刘艳艳, 等. 一种改进 DeeplabV3 网络的烟雾分割算法[J]. 西安电子科技大学学报, 2019, 46(6): 52-59.
- Wang Ziyi, Su Yuting, Liu Yanyan, et al. An improved DeeplabV3 network smoke segmentation algorithm[J]. Journal of Xidian University, 2019, 46(6): 52-59. (in Chinese with English abstract)
- [35] 贺钰博, 刘坤. 基于卷积神经网络的海面显著性目标检测[J/OL]. 计算机工程与应用: [2020-03-08]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20200220.1539>.
- He Yubo, Liu Kun. Sea surface salient target detection based on convolutional neural network[J/OL]. Computer Engineering and Applications: [2020-03-08]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20200220.1539>. (in Chinese with English abstract)
- [36] 王志杰. 基于图像增强的语义分割方法研究与实现[D]. 济南: 山东大学, 2019.
- Wang Zhijie. Research and Implementation of Semantic Segmentation Method based on Image Enhancement[D]. Jinan: Shandong University, 2019. (in Chinese with English abstract)
- [37] Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[C] // International Conference on Computer Vision (ICCV). Venice, Italy, 2017: 2223-2232.
- [38] Liu P, Liu X, Liu M, et al. Building footprint extraction from high-resolution images via spatial residual inception convolutional neural network[J]. Remote Sensing, 2019, 11(7): 830-849.
- [39] 谢亚坤. 基于高分影像阴影的多种场景建筑物高度信息提取方法研究[D]. 成都: 西南交通大学, 2018.
- Xie Yakun. Research on the Height Information Extraction Method of Multiple Scenes Based on High-resolution Image Shadows[D]. Chengdu: Southwest JiaoTong University, 2018. (in Chinese with English abstract)

## Tile edge effect and elimination scheme of image classification using CNN-ISS remote sensing

Duan Zengqiang, Liu Jiedong, Lu Ming, Kong Xiangbin, Yang Na

(College of Land Science and Technology, China Agricultural University, Key Laboratory of Agricultural Land Quality and Monitoring, Ministry of Natural Resources, Beijing 100193, China)

**Abstract:** Semantic segmentation of an image has become a key interdisciplinary application in image processing, computer vision, pattern recognition, and artificial intelligence. In deep learning architectures, the Convolutional Neural Network for Interferometric Semantic Segmentation (CNN-ISS) is widely used in digital image processing and machine vision. The CNN-ISS can be utilized to effectively extract further features, such as texture and geometric features, indicating stronger transfer learning and generalization, compared with traditional image classifications of remote sensing. As such, the CNN-ISS is suitable for the interpretation of high-resolution remote sensing images, identification of complicated features, and crop mapping. In classification, large remote sensing images need to be segmented into specific tiled images, thereby serving as the object of Convolutional Neural Network (CNN) processing. However, artificial image tiling can generate fragments on the edge of a tile, leading to the low classification accuracy of pixels near the edge of the tile. Here, the phenomenon was defined as the edge effect of tiled images, where the classification accuracy of pixels near the edge of the tile was lower than that of the central area. In this work, two indicators were designed, including the error rate with a distance to tile edges (ERD), and the error rate of the whole image (ERW), to quantify the edge effect of CNN-ISS processed tiled images. Meanwhile, the offset positions ( $i, k$ ) were set for the starting point of the shift window to ensure that any pixel on the whole image must be in the central area of the tile generated under a certain offset setting. Then, five technical solutions were obtained to test the minimized edge effect of tiled images using the scores in multiple groups of categories. Taking the Tangshan as the segmented typical rural surface, a DeepLab V3 was selected as the core model of CNN-ISS to analyze the edge effect of the classification. The results showed that the pixel classification accuracy was positively correlated with the distance from the pixel to the edge of a tiled image. The highest error rate was 6.93% occurred along the edge of the tiled image, and the lowest error rate was 3.52% in the center of the tile, indicating the accuracy of the central area was higher than that of the edge. It showed an obvious edge effect of tiled images. In edge effect elimination scheme for the tiled images, the total classification accuracy improved significantly, where the Kappa coefficient and Mean Intersection over Union (mIoU) of the entire image increased 0.012 2 and 1.97 percent point, respectively. Taking the Kappa coefficient, one of the classic accuracy indices for the remote sensing image interpretation, as an example, the order of accuracy including the control group was: solution 2 (0.881 0) > solution 5 (0.878 9) > solution 3 (0.878 8) > solution 4 (0.877 7) > solution 1 (0.875 9) > the control group (0.868 8). Besides, the solutions of edge effects depended mainly on the types of features in the tiled images. The general law was that the tile edge effects of linear features and complex isomers (pit ponds, rural residential areas) were more obviously improved the accuracy, as the solutions were more significantly accurate, compared with that of the base land, or another agricultural land. Compared with the control group, the improvement order of IoU in the solution 2 was: roads (4.13 percent point) > pit ponds (2.97 percent point) > rivers and ditches (1.61 percent point) > rural residential areas (0.65 percent point) > another agricultural land (0.46 percent point). Without changing the core model of CNN semantic segmentation, the elimination scheme for the edge effect of a tile can be used to effectively improve the accuracy of remote sensing image classification, especially for the linear features and complex isomers.

**Keywords:** remote sensing; convolutional neural network; semantic segmentation; image classification; edge effect of tiled images