



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: 智能优化学习的高空间分辨率遥感影像语义分割  
作者: 邵振峰, 孙悦鸣, 席江波, 李岩  
DOI: 10.13203/j.whugis20200640  
网络首发日期: 2021-09-23  
引用格式: 邵振峰, 孙悦鸣, 席江波, 李岩. 智能优化学习的高空间分辨率遥感影像语义分割[J/OL]. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20200640>



**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、外文字符、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.whugis20200640

# 智能优化学习的高空间分辨率遥感影像语义分割

邵振峰<sup>1</sup> 孙悦鸣<sup>1</sup> 席江波<sup>2</sup> 李岩<sup>3</sup>

1 武汉大学测绘遥感信息工程国家重点实验室, 湖北武汉, 430079

2 长安大学地质工程与测绘学院, 陕西西安, 710054

3 内蒙古电子信息职业技术学院, 内蒙古呼和浩特, 010070

**摘要:** 高空间分辨率(高分)遥感影像正被广泛应用, 而传统分类算法在高分遥感影像上表现的精度和效率都较差, 深度学习语义分割算法在高分遥感影像的实际分类中泛化性较差。为了适应大范围高分遥感影像的特点, 提出了一种基于U-Net网络的模拟退火超参数优化与深度可分离卷积语义分割模型。首先在U-Net网络基础上使用了深度可分离卷积模块来进行特征提取, 在保持高效性的同时减少模型的参数量和计算量, 然后利用基于模拟退火的智能优化学习模型搜索网络超参数的全局最优解, 自动优化网络训练初始点。最后在ISPRS2D和GID(Gaofenimage dataset)数据集上进行实验, 与对比实验相比, 在ISPRS2D数据集的分类结果中建筑物、低植被和汽车及总体分类上精度均有提高, 在GID数据集的分类结果中水域、草地、森林及总体分类上精度均有大幅提升, 实验结果验证了本文提出模型的高效性、高精度性与鲁棒性。

**关键词:** 语义分割; 高分辨率遥感影像; 深度学习; 超参数优化

中图分类号: P237 文献标志码: A



**项目资助:** 国家自然科学基金(41890820, 41771452, 41771454, 61806022); 内蒙古自治区自然科学基金(2019MS04017); 内蒙古自治区高等学校科学研究项目(NJZY20277); 地理信息工程国家重点实验室(SKLGIE2018-M-3-4); 长安大学中央高校基本科研业务费专项资金(300102269103, 300102269304, 300102269205)。

**第一作者:** 邵振峰, 博士, 教授。shaozhenfeng@whu.edu.cn

**通讯作者:** 孙悦鸣, 硕士。sunnyueming@whu.edu.cn

席江波, 博士, 讲师。xijiangbo@chd.edu.cn

目前,遥感传感器硬件性能的不断提升,获得的高空间分辨率(高分)遥感影像具有丰富的几何细节、纹理特征和复杂的上下文特征。高分遥感影像<sup>[1]</sup>正在被逐渐应用于各个领域,如国土资源调查<sup>[2]</sup>、城市建设与规划<sup>[3,4]</sup>、农业生产<sup>[5]</sup>、自然灾害和危机管理<sup>[6]</sup>等领域。而遥感影像的语义分割在遥感影像解译中起着重要的作用。本文所说的遥感影像的语义分割与遥感影像的分类是一致的。传统的遥感影像分类方法主要是基于色彩、纹理信息及类似的低级特征来完成。随机森林法是一种使用多棵树来训练和预测样本的分类器<sup>[7]</sup>。马尔可夫随机场法使用无向图模型分割每个像素<sup>[8]</sup>,支持向量机构建超平面来决策分类<sup>[9]</sup>。然而,这些传统的图像语义分割方法针对高分遥感影像精度低,计算效率差,难以满足现代自动化操作的需求。

在过去几年里,深度学习发展迅猛。与传统机器学习的算法相比,深度学习可以提取隐藏在影像的深度抽象特征,且在复杂场景上具有更好的表现。所以,利用深度学习实现高分遥感影像的语义分割为遥感影像分类算法开辟了一条新途径。例如,门计林等人<sup>[10]</sup>提出了一种多结构卷积神经网络特征级联模型在遥感影像上进行土地分类,但该方法运行效率低且分类时个别类别精度反而下降。郑益勤等人<sup>[11]</sup>利用深度信念网络在气象卫星影像上进行强对流云团的提取,但准确度较低,分割边缘较差。季顺平等<sup>[12]</sup>提出一种用于分类的全空洞卷积神经元网络来进行城市土地覆盖分类与变化检测,但精度改善程度不大。陈鹏等<sup>[13]</sup>利用 ResNet 和深度迁移学习来进行建筑物倒塌评估,但对细微特征的提取能力较弱。

而在遥感语义分割的实际应用中,若是人工的为每一个新的数据集去设计和构建一个新的深度学习网络模型,可能会花费大量的时间且需要做一些繁琐的工作。所以,自动机器学习(Automated Machine Learning, AutoML)就成为了深度学习算法从理论到实际应用的关键步骤。其目标是在没有人工干预的情况下获得一个具有良好泛化能力的模型,实现机器学习过程的自动化。而超参数优化(Hyperparameter Optimization,

HPO)则是自动机器学习中重要的一环。我们称定义模型属性或者定义训练过程的参数为超参数。由于深度学习结构的限制,大规模的深度学习网络模型的超参数众多,而各个超参数对深度学习模型的影响并不是独立的,所以超参数的选择对模型最终的效果有极大的影响<sup>[14]</sup>。在一个数据集上表现良好的网络模型可能并不适合于另一个不同的数据集。而在一般情况下,超参数的选择取决于以前的经验和实验结果的组合,这种做法重复实验的效率低且结论具有随机性,不完整。而传统的网格搜索、随机搜索等自动调参算法很难得到全局最优解,且计算量大,容易组合爆炸。因此,非常有必要利用智能优化学习方法实现深度学习网络的超参数优化。

针对以上问题,考虑到实际应用场景中的时间、硬件条件的限制,并针对高分遥感影像数据海量且细节丰富的特征,本文设计了一种基于U-Net网络的模拟退火超参数优化与深度可分离卷积语义分割模型,该模型在保持高效性的同时减少模型的参数量和计算量并实现自动优化网络训练初始点,实现遥感影像地物的精细分类。

## 1 研究方法

本文设计了一种基于U-Net网络的模拟退火超参数优化与深度可分离卷积语义分割模型,并基于该模型实现了高空间分辨率遥感影像的精细分类算法,流程如图1所示。

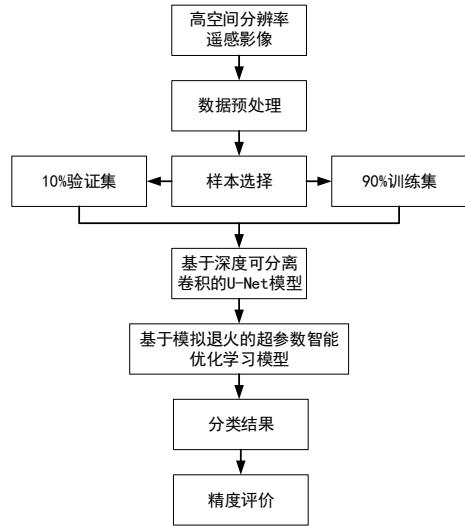


图1 本文算法流程

Fig. 1 Flowchart of the Algorith

基于U-Net网络的模拟退火超参数优化与深度可分离卷积语义分割模型包含编码器和解码器两部分，在编码器中，使用了深度可分离卷积模块来进行特征提取，代替原始U-Net中的标准卷积模块，建立小而高效的语义分割网络模型。并基于模拟退火进行超参数智能优化学习，模拟退火算法可以不断跳出目前的局部最优解来搜索全局最优解，机器自动优化网络模型训练初始点，从而实现高空间分辨率遥感影像的精细分类。

### 1.1 基于 U-Net 网络的模拟退火超参数优化与深度可分离卷积语义分割模型

U-Net基本上可以视为一个对称的结构，它的结构简单清晰，由一个下采样过程和一个上采样过程组成。且U-Net在上采样过程中不断的融入下采样的最终或中间的特征图，融合了多个尺度的特征，深层的卷积仍能够包含浅层的简单特征，使得输出结果更能反映原影像的信息。

然而，在处理高分遥感影像的实际应用

中，庞大而复杂的模型是难以被成功应用的，由于一定的硬件限制因素，常常面临着内存不足、功耗高、延迟高的问题，且它的响应速度慢，很难满足一定场景的实时要求。所以，小而高效的卷积神经网络模型的研究就显得非常必要。研究要减小模型的大小并提升模型的训练和测试速度，且这一切都以保持模型的性能精度为前提。

在本文中，使用了深度可分离卷积模块来进行特征提取，代替原始U-Net中的标准卷积模块，如图2所示为本文中构建的深度可分离卷积模块和基于深度可分离卷积的U-Net网络模型。深度可分离卷积由标准卷积分解而来，先用深度为一的逐通道卷积核分别在不同的输入通道进行滑动，每一次生成一个输出通道，当全部完成后再利用 $1 \times 1$ 的点卷积将各个通道进行跨通道处理调整深度。深度可分离卷积一方面减少了模型的参数量和计算量同时仍保持着较好的性能，它也是MobileNet等<sup>[15]</sup>网络的基础组成部分。

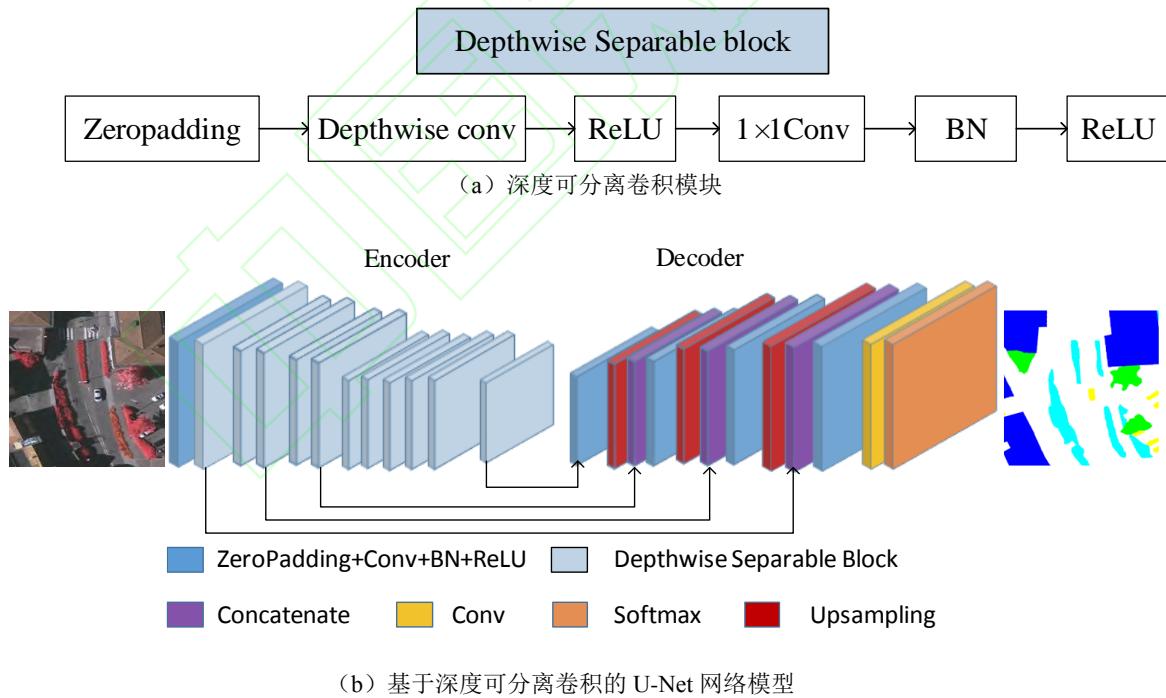


图 2 Depthwise Separable Block and the U-Net Model Based on Depth wise Separable Block

在构建的深度可分离卷积U-Net语义分割模型上使用模拟退火的方法结合遥感数据集优化网络的超参数设置，提高网络性能，具体的模拟退火超参数寻优流程如图3所示，

其中得到的解  $w$  即为需要寻优的超参数组合，目标函数值则为当前  $w$  下模型训练后验证集的精度。

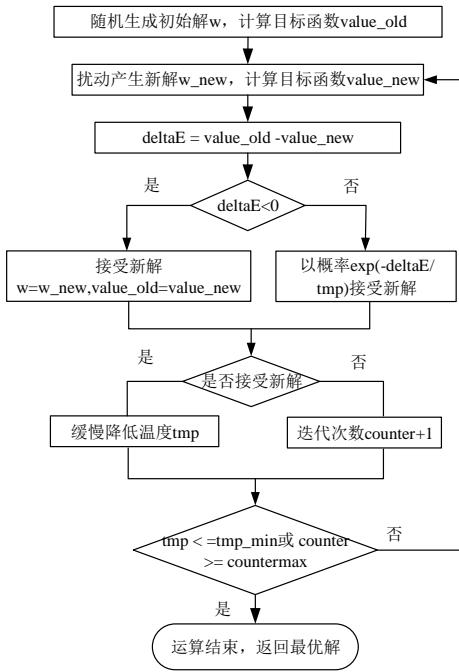


图 3 模拟退火超参数寻优流程图

Fig. 3 Flow Chart of Hyper parameter

#### Optimization in of Simulated Annealing

模拟退火算法<sup>[16]</sup>的思想与晶体冷却相似,若在其冷却过程中温度下降的太快,其中的原子就没有充分的时间来形成晶体的结构,这就会导致形成非晶体即局部最优解。而若我们在此过程中向其中添加一些能量,即以一定的概率去接受较差解,就打破了形成的非晶体即跳出了现在的局部最优解,然后再次进行冷却,重复这个过程,温度缓慢下降,即接受较差解的概率慢慢下降,求解过程逐渐趋于稳定,就能形成有规则晶体即全局最优解,即退火成功。因此可以看出,它实际上是一种模拟晶体冷却的变邻域搜索算法,可以得到全局最优解。

选择的基于语义分割模型需要优化的超参数有:批尺寸(batch size)、时期(epoch)、初始学习率。

##### (1) 批尺寸 (batch size)

指每一次训练时输入网络模型的样本量的大小,批尺寸太少容易陷入局部最优,一般将其增大可加快网络的收敛,提高训练速率,但硬件条件有限,过大的批尺寸会导致内存不足或程序内核崩溃,且得到相同训练精度所需要的迭代次数更大。

##### (2) 时期 (epoch)

指所有训练集数据前向传播和后向传播后更新参数的过程,一般来说 epoch 越大得到的精度越高,但 epoch 过大容易过拟合,且会导致训练时间过长,特别在训练后期,随着 epoch 增加而精度的涨幅微小,大大浪费训练时间。

#### (3) 初始学习率

学习率决定着网络权重的更新速度,过大会造成模型无法收敛;过小会造成模型收敛速度的下降或陷入局部最优解,不同的优化算法决定不同的学习率。

#### 1.2 精度评价

本文使用的评价标准为总体精度  $O$ 、F1 分数和平均交并比 (mean inter section over union, MIoU)。其表达式分别为:

$$O = \frac{TP+TN}{TP+FN+FP+TN} \quad (1)$$

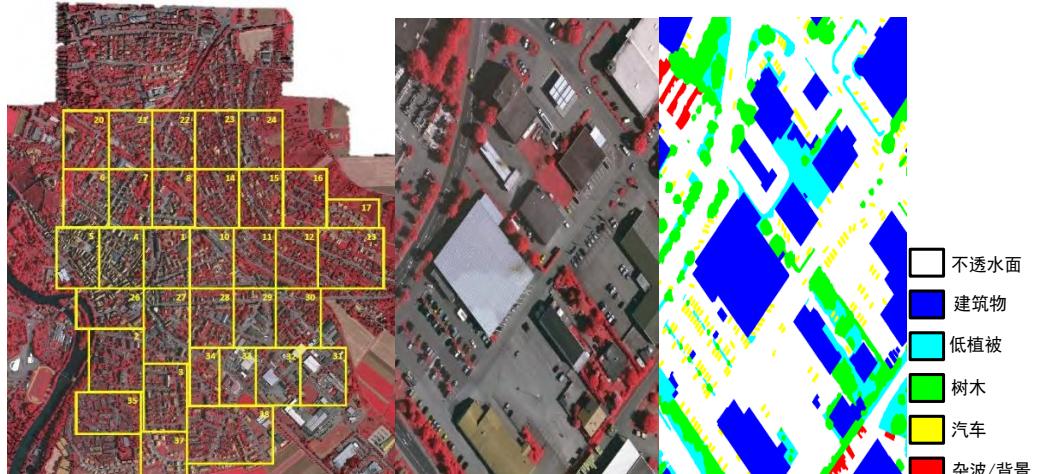
$$F1 = 2 \frac{P \cdot R}{P+R} \quad (2)$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{TP+FN+FP} \quad (3)$$

式中,精确率  $P = \frac{TP}{TP+FP}$ , 召回率  $R = \frac{TP}{TP+FN}$ ; 真阳性(TP)表示预测为真,实际是真;假阳性(FP)表示预测为真,实际是假;假阴性(FN)表示预测为假,实际是真;真阴性(TN)表示预测为假,实际是假。

## 2 实验与分析

在本文实验中使用的第一组数据集是 ISPRS2D 语义分割挑战赛的公开数据集,它是德国 Vaihingen 地区的航空影像数据。如下图 4 所示,此数据集中包含了 33 幅不同大小高分辨率真正射(Top)遥感影像,分辨率为 9 cm。其中 16 幅影像包含地面真实标签图,被分为六类:不透水面,建筑物,低植被,树木,汽车,杂波/背景。第二个数据集是 GID(Gaofen Image Dataset)<sup>[17]</sup>,这是一个用于像素级分类的高分辨率遥感影像数据集。它包含中国 60 多个不同城市采集的 150 幅高分辨率高分 2 号 (GF-2) 的图像。包含五个类别:建筑、农田、森林、草地和水域。其余部分用黑色标记。图 5 为该数据集的示意图。



(a) 33幅影像贴片图 (b) 贴片32的正射影像图和地表真实标签图

图 4ISPRS2D 数据集示意图

Fig. 4Schematic Diagram of the ISPRS2D Dataset

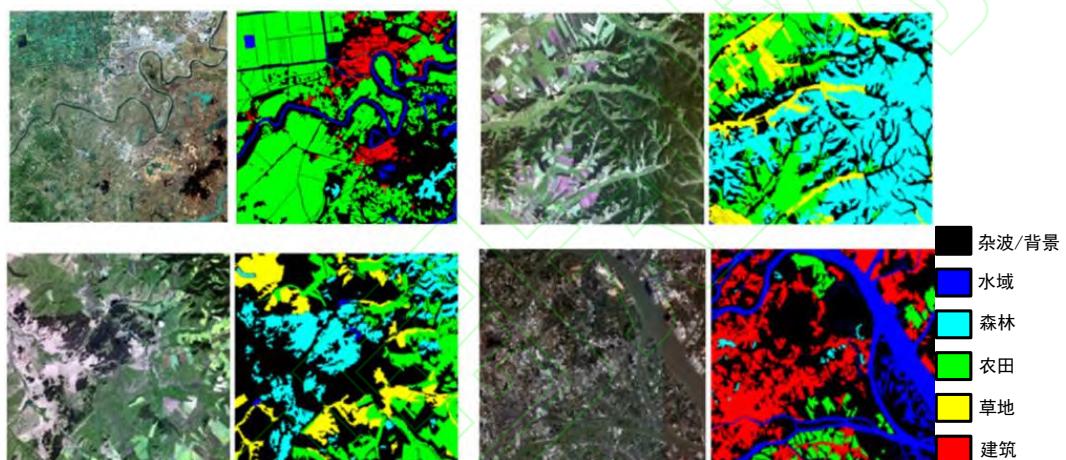


图 5GID 数据集示意图

Fig. 5Schematic Diagram of the GID Dataset

本文实验的环境为 Windows10.64 位操作系统，并在 Python3.5 下运行并配合开源深度学习框架 Tensorflow1.13.2 和 keras2.1.5，处理器为 IntelCorei7-8700KCPU3.7GHz，内存为 32GB，配备 GeForceGTX1080Ti。

## 2.1 基于 U-Net 网络的模拟退火超参数优化与深度可分离卷积语义分割模型实验

由于训练集图片较少且尺寸过大，需要对数据集进行进一步处理，在 ISPRS2D 数据集上使用有真实标签的 15 张影像进行训练和验证，剩余 1 张  $1388 \times 2555$  影像进行测试。在 GID 数据集上随机选取 100 张有真实标签的影像进行训练和验证，1 张  $6800 \times 7200$  影像进行测试。分别对两个训练集中的遥感影像数据先进行缩放、翻转、裁剪等数

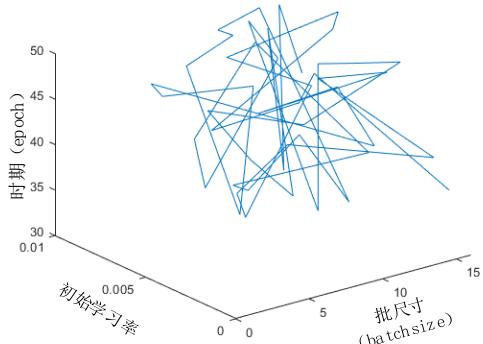
据增强方法后进行规定尺寸为  $416 \times 416$  的随机裁剪。两个扩充后的训练数据集分别为 1000 张，其中进行模型训练的图片有 900 张，验证的图片有 100 张。

模型使用 Adam 优化器进行梯度下降，当验证集的 loss 值三次不下降的话，学习率减半。模拟退火的参数初始化设置为：初始温度  $tmp=10$ ，退火概率  $alpha=0.89$ ，最低温度  $tmp\_min=0.001$ ，最大迭代次数  $counters\_max=100$ 。

超参数优化范围设置为：batchsize 通常取值为 [2,32]，在本次实验中，由于内存资源有限，batchsize 的优化范围为 [2,16]。为保证任意一组超参数组合都能使模型得到收敛，epoch 的优化范围为 [30,50]。在本实

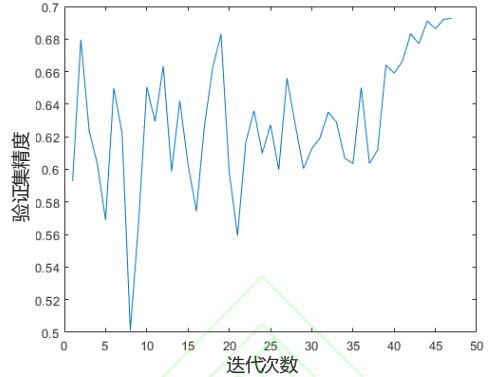
验中语义分割网络使用的是 Adam 优化器，根据通常学习率的初始值，在本次实验中设定的优化范围为[0.01,0.001]。

如下图 6 所示为在 ISPRS2D 数据集上



(a) 模拟退火过程的线状图

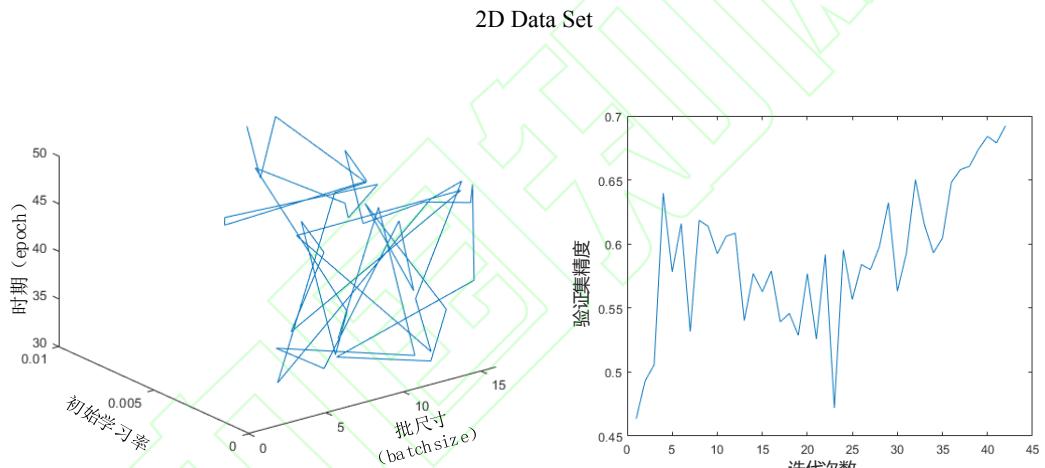
的基于模拟退火的超参数寻优过程及相应的验证集精度的变化过程。图 7 所示则为在 GID 数据集上的基于模拟退火的超参数寻优过程及相应的验证集精度的变化过程。



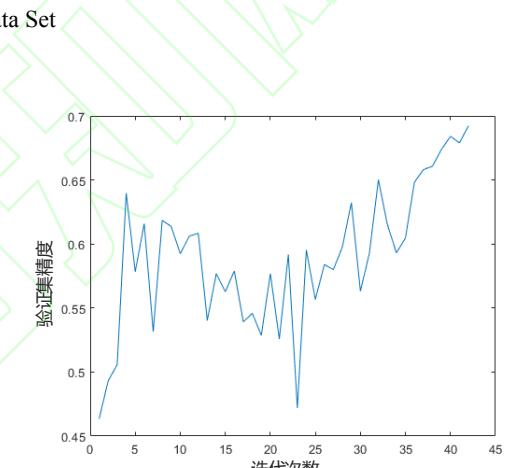
(b) 模拟退火过程 val\_acc 的变化图

图 6 ISPRS 2D 模拟退火算法寻找最优超参数的过程图

Fig.6 Process Chart of Simulated Annealing Algorithm for Finding the Optimal Hyper parameters of ISPRS



(a) 模拟退火过程的线状图



(b) 模拟退火过程 val\_acc 的变化图

图 7 GID 模拟退火算法寻找最优超参数的过程图

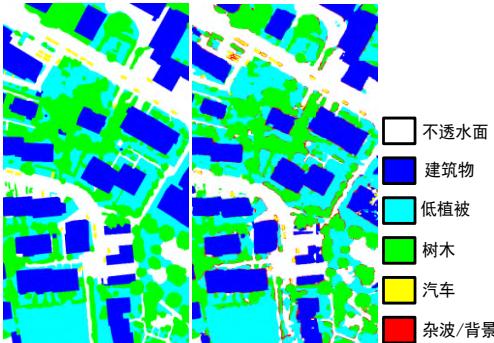
Fig.7 Process Chart of Simulated Annealing Algorithm for Finding the Optimal Hyper parameters of GID Data Set

由上述图表可以看出，在超参数搜索过程中 ISPRS2D 验证集的精度变化范围为[0.559-0.693]，GID 验证集的精度变化范围为[0.463-0.693]，可以看出超参数的选择对于网络模型的训练精度影响是非常大的，且在大型深度学习网络中，超参数的可选个数和范围更大且超参数之间相互组合的结果无法预料，人工进行合适的超参数选择的可能性微乎其微，非常需要智能优化学习的方法来使计算机自动进行超参数的优化，且由上图所示模拟退火算法不断跳出局部最优值，去寻找全局最优值，提升了超参数优化

的搜索精度。

最终得到在 ISPRS2D 数据集上的最优超参数组合：batchsize 为 13，epoch 为 45，初始学习率为 0.006916985153914473。得到 GID 数据集上的最优超参数组合：batchsize 为 11，epoch 为 47，初始学习率为 0.0021372187899398104。并将上述智能优化学习后搜索到的最优超参数组合带入到训练数据集中进行实验，在 ISPRS2D 数据集上得到分类结果如下图 8、9 所示，并在表 1 中与 SVL\_1<sup>[18]</sup>，SegNet<sup>[19]</sup>，FCN-8s<sup>[19]</sup>，U-Net+ $CE_{loss}$ <sup>[20]</sup>方法进行对比；在 GID 数据

集上得到分类结果如下图 10 和 11 所示并在下表 2 中与 Deeplab<sup>[21]</sup>、Pix2pix<sup>[21]</sup>、FCN<sup>[21]</sup>、CRFAS<sup>[21]</sup>方法进行对比。



(a) 真实标签图 (b) 本文模型的预测结果

图 8 ISPRS 2D 的真实标签图和本文模型的语义分割结果

Fig. 8 Real Label Map and Semantic

表 1 ISPRS 2D 数据集上本文方法与其他方法对比表

Tab.1 Comparison Table of Our Method and Other Methods of ISPRS 2D Data Set

模型方法	不透	建	低	树	汽	总体
	水面	筑物	植被	植被	车	精度
SVL_1 <sup>[18]</sup>	86.3	90.8	78.2	84.2	56.8	84.7
SegNet <sup>[19]</sup>	<b>87.8</b>	91.6	74.0	84.5	76.9	85.0
FCN-8s <sup>[19]</sup>	87.3	90.3	73.7	84.9	68.8	84.7
U-Net+ $CE_{loss}$ <sup>[20]</sup>	85.82	90.5	73.6	83.3	67.2	80.1
<b>本文实验</b>	84.8	<b>93.8</b>	<b>81.5</b>	<b>85.1</b>	<b>81.4</b>	<b>86.5</b>

在表 1 中, 各个类别的语义分割精度指标为 F1 分数 (%), 总体精度指标为总体分类精度 O (%). 加粗字体为各类最高语义分割精度。由上述图表可知, 在 ISPRS2D 数据集的分类结果中, 建筑物的综合分割精度最高, 且可以准确分割建筑物的大体轮廓; 由于汽车的尺度较小给准确分割带来了一定难度, 且其外在有一定的多样性, 且存在阴影部分, 容易与路面发生混淆, 很多属于汽车的像素被错分为不透水面这一类, 或被误认为是背景; 且由于低植被和树这两类的光谱特征较为相似, 这两类的混淆问题比较严重; 相较于其他分类方法, 本文方法在建筑物, 低植被, 树、汽车和总体精度上都取得了最高的分类精度, 分割的总精度达到了 86.5%, 可以说分类效果已经较为理想。

Segmentation Results of ISPRS 2D Data Set

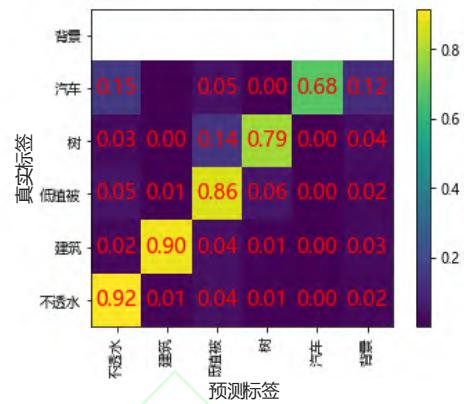
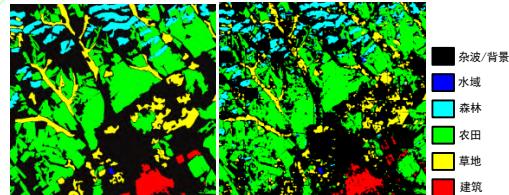


图 9 ISPRS 2D 实验结果的混淆矩阵

Fig. 9 Confusion Matrix of Results of ISPRS 2D Data

Set

Tab.1 Comparison Table of Our Method and Other Methods of ISPRS 2D Data Set



(a) 真实标签图 (b) 本文模型的预测结果

图 10 GID 的真实标签图和本文模型语义分割结果

Fig. 10 Real Label Map and Semantic

Segmentation Results of GID Data Set

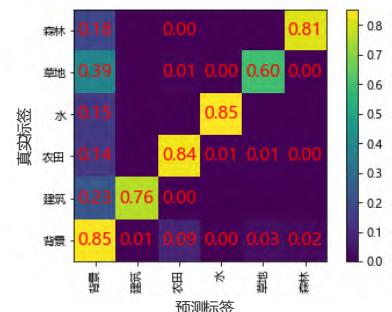


图 11 GID 实验结果的混淆矩阵

Fig. 11 Confusion Matrix of Results of GID

表 2 GID 数据集上本文方法与其他方法对比表

Tab.2 Comparison Table of Our Method and Other Methods of GID Dataset

模型方法	农田	水域	建筑	草地	森林	MIoU
Deeplab <sup>[21]</sup>	<b>87.9</b>	7.1	<b>98.9</b>	6.2	79.9	49.9
Pix2pix <sup>[21]</sup>	81.0	0	92.2	0	62.4	39.8
FCN <sup>[21]</sup>	84.8	34.1	96.6	1.2	76.2	49.9
CRFAS <sup>[21]</sup>	85.7	10.0	98.5	18.5	81.9	51.4
<b>本文实验</b>	86.6	<b>91.4</b>	85.9	<b>73.2</b>	<b>88.5</b>	<b>62.2</b>

在表 2 中,各个类别的语义分割精度指标为 F1 分数 (%),加粗字体为各类最高语义分割精度。由上述图表可知,在 GID 数据集的分类结果中,本文方法也获得了最高的平均精度。从表 2 可以看出,该方法的平均准确度分别比 FCN、Deeplab 高 12.3%,比 pix2pix 和 CRFAS 高 22.4% 和 10.8%。对于水域的综合分割精度最高,且可以准确分割建筑物的大体轮廓;但由于草地的形状不规则,边界比较模糊,容易与背景发生混淆,很多属于草地的像素被错分成背景这一类;本文网络模型采用跳跃连接融合浅层特征与深层特征,浅层特征提高小目标分割能力,深层特征可预测分割大目标,相比其他方法,本文提出的方法对于水域,草地类的分类精度有显著提高。

对于在 ISPRS2D 和 GID 数据集上的语义分割结果可以看出,基于 U-Net 网络的模拟退火超参数优化与深度可分离卷积语义分割模型的高分辨率影像语义分割方法相较于其他方法,在大部分类别和总体的分割精度上都有了一定的提高,充分肯定了模型的可行性和可靠性。

### 3 讨论与结论

基于本文的试验结果分析可得:

(1) 在语义分割网络中利用深度可分离卷积模块来进行特征提取,可在保持精确度和高效性的同时减少模型的参数量和计算量。

(2) 在超参数搜索过程中验证集的精度变化范围很大,证明超参数的选择对于网络模型的训练精度影响非常大。利用模拟退火算法进行网络模型超参数的智能优化学

Dataset

习十分有效可靠,可实现自动优化网络训练初始点,优化后影像分类的各个类别和总体的分类精度都有了一定的提高。

(3) 在高空间分辨率遥感影像中小目标对象的提取和不规则对象提取边缘的优化方面仍有改进的空间。

本文提出了一种基于 U-Net 网络的模拟退火超参数优化与深度可分离卷积语义分割模型,使用深度可分离卷积 U-Net 的网络复杂度小,实现了高效语义分割。本文基于模拟退火实现了超参数优化,优化了网络训练过程,提升了语义分割模型泛化性能。实验结果表明,模拟退火超参数优化使得各个类别和总体的语义分割精度都有了一定的提高,且与其它语义分割模型结果精度相比,该方法能够得到更好的语义分割效果,肯定了基于 U-Net 网络的模拟退火超参数优化与深度可分离卷积语义分割模型的可行性和可靠性。

### 参 考 文 献 (References)

- [1] Chen Ling,Jia Jia,Wang Haiqing. An overview of applying high resolution remote sensing to land and resources survey[J]. Remote Sensing for Land and Resources,2019, 31 (1):1-7(陈玲,贾佳,王海庆.高分遥感在自然资源调查中的应用综述[J].国土资源遥感,2019,31 (1) :1 -7.)
- [2]Zhang Hanchao,Ning Xiaogang,Wang Hao,et al.High accuracy urban expansion monitoring and analysis of China's provincial capitals from 2000 to 2015 based on high-resolution remote sensing imagery[J].Acta Geographica Sinica, 2018,73(12):2345-2363(张翰超,宁晓刚,王浩,等.基于高分辨率遥感影像的2000-2015年中

- 国省会城市高精度扩张监测与分析[J].地理学报,2018,73(12):2345-2363.)
- [3]Yuan Jingwen, Wu Chen, Du Bo, et al. 2020. Analysis of landscape pattern on urban land use based on GF-5 hyperspectral data. Journal of Remote Sensing (Chinese),24 (4): 465-478(袁静文,武辰,杜博,等.高分五号高光谱遥感影像的城市土地利用景观格局分析 [J].遥感学报,2020,24(04):465-478.)
- [4]PanYin,Shao Zhenfeng,Cheng Tao,et al.Analysis of Urban Water logging Influence Based on Deep Learning Model[J].Geomatics and Information Science of Wuhan University, 2019,44(1):132-138.(潘银,邵振峰,程涛,贺蔚.利用深度学习模型进行城市内涝影响分析[J].武汉大学学报·信息科学版,2019,44(1):132-138.)
- [5]Yang Zeyu,Zhang Hongyan,Ming Jin,et al.Extraction of Winter Rapeseed from High-Resolution Remote Sensing Imagery via Deep Learning[J].Bull Surv Map,2020(9):110-113. (杨泽宇,张洪艳,明金,冷伟,刘海启,游炯.深度学习在高分辨率遥感影像冬油菜提取中的应用 [J].测绘通报,2020(9):110-113.)
- [6]Tang Yao,Wang Lijuan,Ma Guochao,et al.Disaster Monitoring and Application Prospect Analysis of the Jinsha River Landslide Based on "Gaofen+"[J].Geomatics and Information Science of Wuhan University, 2019,44(7):1082-1092. (唐尧,王立娟,马国超,贾虎军,李洪梁,尹恒.基于“高分+”的金沙江滑坡灾情监测与应用前景分析 [J].武汉大学学报·信息科学版,2019,44(7):1082-1092.)
- [7]Zhu Wei.Semantic Segmentation of Street Scene Based on Random Forest Algorithm[D]. Harbin University of Science and Technology, 2019.(朱卫. 基于随机森林算法的街道场景语义分割 [D]. 哈尔滨理工大学, 2019.)
- [8]Qi Xianghui.Research on Image Segmentation Algorithm Based on MRF and Fuzzy Clustering[D].Lanzhou University of Technology,2019(亓祥惠. 基于MRF与模糊聚类的影像分割算法研究[D]. 兰州理工大学, 2019.)
- [9]Bruzzone L,Carlin L. A Multilevel Context-Based System for Classification of Very High Spatial Resolution Images[J]. IEEE Transactions on Geoscience&Remote Sensing, 2006,44(9): 2587-2600.
- [10]Men Jilin,Liu Yueyan,Zhang Bin,et al.Land Use Classification Based Network on Multi-structure Convolution Neural 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]Zheng Yiqin,Yang Xiaofeng, Li Ziwei. Detection of Severe Convective Cloud over Sea Surface from Geostationary Meteorological Satellite Images Based on Deep Learning[J]. Journal of Remote Sensing, 2020,24(1): 97-106.(郑益勤, 杨晓峰, 李紫薇.深度学习模型识别静止卫星图像海上强对流云团[J].遥感学报, 2020,24(1): 97-106.)
- [12]Ji Shunping,Tian Siqi, Zhang Chi.Urban Land Cover Classification and Change Detection Using Fully Iltrous Convolutional Veural Vetwork[J].Geomatics and Information Science of Wuhan University, 2020, 45(2):233- 241.(季顺平, 田思琦, 张驰.利用全空洞卷积神经元网络进行城市土地覆盖分类与变化检测[J].武汉大学学报·信息科学版, 2020, 45(2):233-241.)
- [13]Chen Peng, Wang Benkang, Gao Sa, et al.Building Collapse Assessment with Residual Network[J].Geomatics and Information Science of Wuhan University,2020,45(8):1179-1184.(陈鹏, 汪木康, 高飒, 等.利用ResNet进行建筑物倒塌评估[J].武汉大学学报·信息科学版, 2020 , 45(8):1179-1184. )
- [14]Breuel T M. The Effects of Hyperparameters on SGD Training of Neural Networks[J]. arXiv preprint arXiv:1508.02788, 2015.
- [15]Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient Convolutional Neural Networks for

- Mobile Vision Applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [16]Kang Z, Qu Z. Application of BP Neural Network Optimized by Genetic Simulated Annealing Algorithm to Prediction of Air Quality Index in Lanzhou[C].2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA). IEEE, 2017: 155-160.
- [17]Tong XY, Xia GS, Lu Q,et al. Land-Cover Classification with High-Resolution Remote Sensing Images Using Transferable Deep Models[J]. Remote Sens. Environ. 2020, 237, 111322.
- [18]Gerke M. Use of the Stair Vision Library within the ISPRS 2D Semantic Labeling Benchmark(Vaihingen).Holland:University of Twente.2014.
- [19]Pan Xuran, Gao Lianru, Marinoni A, et al. Semantic Labeling of High Resolution Aerial Imagery and LiDAR Data with Fine Segmentation Network[J]. Remote Sensing, 2018, 10(5):743.
- [20]Dong R, Pan X, Li F. DenseU-Net-based Semantic Segmentation of Small Objects in Urban Remote Sensing Images[J]. IEEE Access, 2019,(99):1.
- [21]Xiong D , He C , Liu X , et al. An End-To-End Bayesian Segmentation Network Based on a Generative Adversarial Network for Remote Sensing Images[J]. Remote Sensing, 2020, 12(2):216.

## Intelligent Optimization Learning for Semantic Segmentation of High Spatial Resolution Image

SHAO Zhenfeng<sup>1</sup> SUN Yueming<sup>1</sup> XI Jiangbo<sup>2</sup> LI Yan<sup>3</sup>

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079,China

2 College of Geology Engineering and Surveying, Chang'an University, Xi'an 710054,China

3 Inner Mongolia Electronic Information Vocational Technical College, Huhhot 010070,China

**Abstract:** As the hardware of remote sensing sensors continuously strengthened, the high-spatial resolution remote sensing images are widely used. However, typical classification algorithms have poor accuracy and efficiency in classifying high-resolution remote sensing images, and deep learning semantic segmentation algorithm also has poor generalization in the classification of high-resolution remote sensing images, which make the semantic segmentation of high-resolution images more challenging. In order to adapt to the large-scale high-resolution images, we design a semantic segmentation model of simulated annealing hyper parameter optimization and depth wise separable convolution based on U-Net (Baseline). Firstly, the depth wise separable convolutional module is used to extract the features on the baseline. Then the intelligent optimization learning model based on simulated annealing searches the global optimal solution of the hyper parameters. Finally, experiments were carried out on the ISPRS2D and GID(Gaofen Image Dataset). Compared to other methods, the accuracies of the classification results of the ISPRS2D Dataset for buildings, low vegetations and cars, as well as the general classification were improved based on the proposed method. In the classification results of GID, the accuracies of water, grassland, forest and general classification have been greatly improved by the proposed method. The results of the studies demonstrated the higher efficiency, higher accuracy and robustness of the proposed model.

**Key words:** semantic segmentation; high-resolution image; deep learning; hyper parameter optimization

**Corresponding author:**SUN Yueming, postgraduate. E-mail: sunyueming@whu.edu.cn

XI Jiangbo,PHD, lecturer. E-mail: xijiangbo@chd.edu.cn

Foundation support: The National Natural Science Foundation of China under Grants (41890820,41771452, 41771454 and 61806022);Natural Science Foundation of Inner Mongolia Autonomous Region (2019MS04017);Scientific research project of colleges and universities in Inner Mongolia Autonomous Region (NJZY20277);State Key Laboratory of Geo-Information Engineering(SKLGIE2018-M-3-4);the Fundamental Research Funds for the Central Universities, Chang'an University (300102269103, 300102269304 and 300102269205)

