

边缘增强的 BECU-Net 模型高分辨率 遥感影像耕地提取

董张玉^{1,2,3}, 李金徽¹, 张晋¹, 于金秋¹, 安森¹

1. 合肥工业大学 计算机与信息学院, 合肥 230601;
2. 合肥工业大学 工业安全与应急技术安徽省重点实验室, 合肥 230601;
3. 合肥工业大学 智能互联系统安徽省实验室, 合肥 230601

摘要: 耕地作为国家粮食生产的重要保障, 其空间分布是粮食安全评估、土地资源管理等领域的主要依据。为解决现有的耕地信息提取方法忽视地块的差异化特征和边缘细节蕴含的丰富信息, 且提取结果碎片化、边界模糊问题。本研究以耕地为研究对象, 采用一种结合 EfficientNet 骨干网络和 U 型框架构建的改进型耕地信息提取模型 BECU-Net (Boundary Enhancement Classification U-Net), 并为实现边缘特征和深度特征的信息互补, 设计由 CoT 模块 (Contextual Transformer Module)、门控卷积、scSE (Spatial-Channel Squeeze and Excitation) 注意力机制形成的边缘分支子网络, 来提高模型处理边界信息的专注度。同时, 构建含约束项的联合型边缘增强损失函数 BE-LOSS (Boundary Enhancement Loss) 进一步完善模型运算性能。使用 GID 高分二号 RGB-NIR 四波段数据, 与梯度、指数、纹理特征图共同构建耕地特征机制。并分别与不同网络结构、不同损失函数的模型进行对比。结果表明: 改进算法的总体精度和 F1 分数均有改善, 相比于 DeeplabV3+ 网络, 提取精度提升 2.24%, F1 分数提升 1.77%。本研究提出的新算法可为进一步解决耕地信息提取时边界模糊问题提供技术参考, 为复杂交界的精准划分提供理论支撑。

关键词: 遥感, 边缘增强, 耕地提取, 语义分割, U-Net, 高分影像

中图分类号: TP751.1/P2

引用格式: 董张玉, 李金徽, 张晋, 于金秋, 安森. 2023. 边缘增强的 BECU-Net 模型高分辨率遥感影像耕地提取. 遥感学报, 27(12): 2847-2859

Dong Z Y, Li J H, Zhang J, Yu J Q and An S. 2023. Cultivated land extraction from high-resolution remote sensing images based on BECU-Net model with edge enhancement. National Remote Sensing Bulletin, 27(12): 2847-2859 [DOI:10.11834/jrs.20222268]

1 引言

耕地是人类社会赖以生存的基本资源和农业的根本。坚守耕地红线, 关乎粮食安全、国民经济和可持续发展。然而, 随着社会经济进步, 人类对土地资源进行优化配置以满足生产生活的物质需求日益增长, 居民点、林地、公路等利用类型对耕地空间不同程度的占用, 造成耕地物理界限复杂化和边缘细碎化 (杜国明等, 2021)。因此, 耕地覆盖信息的精准监测, 对于落实农田整治、强化监督、规划管控等政策具有重要意义。

传统的土地覆盖调研工作依赖专家先验知识进行目视解释 (马霭乃, 1987; 杨博和刘义, 2008)。20 世纪 70 至 80 年代以来, 国内外学者针对如何提高遥感影像分类精度、降低人工成本展开大量研究。

早期分类技术的研究主要集中在统计模式识别和基于知识的分类规则上 (甘甫平等, 1999), 使用边缘作为图像分类的重要指标 (Mather, 1986)。为进一步利用图像边缘作为重要分类特征所蕴含的丰富信息、避免分类过程的边缘模糊和错判, (Gong 和 Howarth, 1990) 等将拉普拉斯算子生成

收稿日期: 2022-05-27; 预印本: 2022-09-23

基金项目: 安徽省自然科学基金 (编号: 2108085MF233); 安徽省重点研究与开发计划 (编号: 202004a07020030); 中央高校基本科研业务费专项 (编号: JZ2021HGTB0111)

第一作者简介: 董张玉, 研究方向为多源信息处理与遥感信息获取、处理及其应用。E-mail: dzyhfut@hfut.edu.cn

的边缘密度图像作为结构信息纳入传统分类,使得城市与农村边界划分结果的光谱混乱和空间混淆现象有所改善。(Pan等, 2012)分别采用SCM方法提取硬边, HCM方法提取软边, 独立处理不同类型的边缘信息, 一定程度提高了耕地提取精度。然而, 机器学习分类器(如马氏距离MD、支持向量机SVM、最大似然MLC (Verma等, 2017)、决策树DT (Lee等, 2020; Wu等, 2013)等)多基于光谱特征或空间格局的像素聚合 (Fan等, 2021), 本质上未有效利用图像的上下文信息、形态等深层语义, 从而忽略了高分辨率遥感影像的其他空间特征, 在耕地与多种覆盖类型的复杂交界处难以实现耕地区域的准确划定。

这些分类算法很大程度地依赖手工特征的设计和提取, 一些难以设计或隐含的图像特征往往被弃用。图1为分别使用 (Liu等, 2015) 方法和 (Xu等, 2019) 方法得到的提取结果。(Liu等, 2015) 基于两阶段图像分类方法, 分别使用监督迭代条件模式 ICM 和局部二值模式 LBP 提取耕地信息, 图1 (c) 为使用该方法的提取效果, 显示斑块边缘曲折, 图块漏检较多。(Xu等, 2019) 提出了一种基于分层对象的农田提取方法实现精细化的耕地分割, 图1 (f) 为该方法的提取效果图斑边缘毛糙, 非耕地区域存在噪声点。

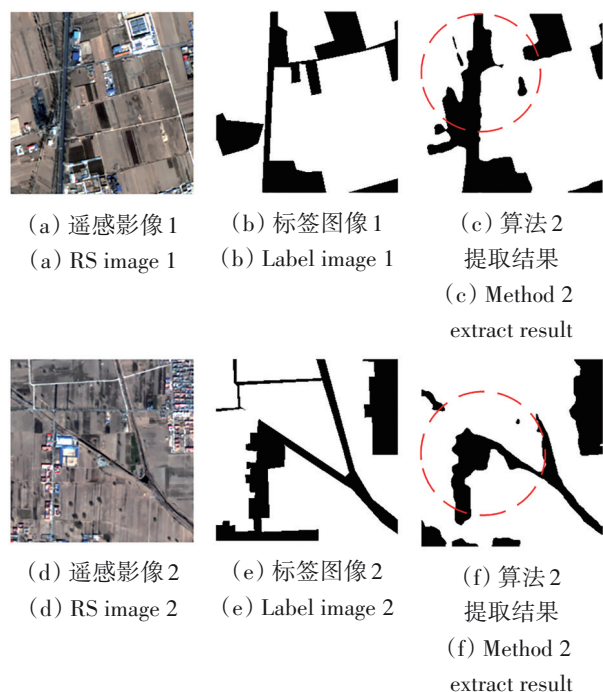


图1 耕地提取时的边缘模糊现象

Fig. 1 Edge ambiguity in cultivated land extraction

21世纪以来, 随着“数字农业”体系不断完善和计算机技术的迅速发展, 人工智能科技广泛应用于国土资源信息化建设。其中, 新兴的深度学习算法通过构建深度神经网络, 由浅至深地提取遥感影像的多层次特征。自2012年 (Krizhevsky等, 2012) 以来, 各类丰富的学习网络和训练数据集为耕地信息的自动提供理论支撑。(He等, 2020) 利用整体嵌套边缘检测 (HED) 网络检测到的边缘信息修正FCN网络 (Long等, 2015) 分割结果, 并在GID数据集 (Tong等, 2020) 测试实验中对各覆盖类型均达到较高的分类精度。(Zheng等, 2022) 提出一种新的多粒度边缘保留优化方法 MRF-MEO, 配合 DeeplabV3+ (Chen等, 2018) 语义分割网络, 提高分割精度, 改善边缘模糊和过度平滑的问题。但 DeeplabV3+ 网络训练和推理速度慢、批次量大, 且仅将边缘增强作为深度学习的后处理步骤的设计思路, 增强的边缘特征并不参与高级语义特征的监督和损失反传, 优化效果有限。

编码—解码式语义分割网络U-Net (Ronneberger等, 2015) 及其各类衍生网络 (Zhou等, 2020; Bao等, 2021; Xu等, 2020) 在全卷积FCN的原理基础上, 作出了更适应二值分类任务的调整, 巧妙设计跳跃连接保留低级语义, 并利用完全对称的网络结构使分割图恢复出更加精细的边缘信息, 研究表明 UNet 系列结构在耕地信息的有效提取上表现出巨大潜力。(Xia等, 2018) 分别使用边缘检测模型 RCF 和语义分割模型 UNet 准确获取常规块状耕地的软硬边缘属性信息, 此类方法两模型间没有实现边缘信息和语义特征的互通互补, 因而对于边缘复杂、居民点密集、地块狭小的区域异质性遥感图像提取效果不佳。综上, 为进一步实现耕地信息的准确提取, 尤其是复杂地形条件下的边界判断, 亟待解决以下技术问题: (1) 如何使边缘信息和语义特征间有效地互通互补, 进行耕地地块的精准覆膜; (2) 如何针对耕地信息提取和边缘增强设计更高效的损失函数, 使网络高效率地随着一次次的反向传播不断修正训练技术存在的问题, 本文提出了一种基于高分辨率遥感影像的耕地信息提取方法, 该方法耦合了语义分割模型和边缘增强模型, 利用增强边界类和聚焦训练的方法来强化训练样本; 设计与网络功能更适应的边缘损失函数, 进一步提升训练精度和模型性能, 以期对现有的耕地信息提取方法中存

在的边缘模糊问题提供理论支撑。

2 改进的 BECU-Net 模型

2.1 多特征输入层

耕地斑块较其他地表类型具有明显差异性特征，在数据输入网络之前，本文对预处理之后的数据进行特征预提取，将预提取得到的特征作为网络输入的补充图层，包括 2 个指数特征图层和 3 个纹理特征图层。输入结构的调整，提高了网络的特征表达能力，使语义分割网络在学习过程中能够更有效地学习到这些深层次的特征，优化了网络的学习结构。

2.1.1 指数特征

为全面考虑在耕和休耕期的作物、土地差异情况，本文从 2 个角度设计指数特征预提取方案，分别是改良土壤调整植被指数 MSAVI (Modified Soil Adjusted Vegetation Index) (Wu 等, 2019) 和绿—蓝光不透水面指数 (Green-Blue Impervious Surface Index, GBISI) (杨栩等, 2020)。其中，改良土壤调整植被指数 MSAVI 相比于 NDVI 能够减轻土壤对作物监测结果的影响，对于像耕地这样裸土比例较高的土地类型更加适用，其计算公式如下：

$$MSAVI = \frac{2 \times NIR + 1 - \sqrt{(2 \times NIR + 1)^2 - 8 \times (NIR - RED)}}{2} \quad (1)$$

式中，NIR 为近红外波段的反射率，RED 为可见红光波段的反射值。

绿—蓝光不透水面指数 GBISI 是一种使用可见光计算影像不透水面的算法，可以快速区分耕地和人造地面，同时使用可见光数据计算指数，解决了高分数据缺乏中红外、热红外波段的局限，计算公式如下：

$$GBISI = \frac{k\rho_{blue} - \rho_{green} + l}{\sqrt{k^2 + 1}} \quad (2)$$

式中， k 和 l 根据研究区数据进行回归分析，在本文中， $k = 0.74$ ， $l = 0.67$ ； ρ_{blue} 和 ρ_{green} 分别表示各蓝色和绿色可见光波段像元值。

2.1.2 纹理特征

纹理分析广泛应用于遥感分类，灰度共生矩阵 GLCM (Gray Level Concurrence Matrix) 算法是最常见和最广泛接受的统计纹理分析方法之一。GLCM 是 2 个相邻像素之间方向和距离关系的函数，

可以定量反映耕地图斑中的纹理分布情况。

在本文中，使用 GLCM 算法计算样本数据的多种纹理特征 (Chen 等, 2016)。随后使用随机森林多次重复选取验证，得到熵、同质性、对比度 3 个显著特征类型。

2.2 BECU-Net 网络整体结构

基于 UNet 架构的模型被广泛应用于遥感图像分析，其在对称的编解码结构中提出“跳跃连接”，将解码输出与各级解码器进行特征级联，融合了底层特征的同时重建空间信息。本文在 UNet (Ronneberger 等, 2015) 基础上，提出一种级联边缘增强分支的 BECU-Net (Boundary Enhancement Classification U-Net) 耕地信息提取网络，其网络整体结构如图 2 所示，在 U 型架构之上使用多分支编码结构，整体网络包括深度特征分支、边缘增强分支和解码器 3 部分。如图，深度特征分支网络采用 EfficientNet-B7 (Tan 和 Le, 2020) 作为的基本架构，模型自动缩放和扩展网络长度、宽度和分辨率，且不改变基线网络中预定义的原始架构，从而获得更好的模型性能。边缘增强分支网络具体设计了 2 个模块 (包括 CoT 模块、门控卷积) 在 UNet 的基础上尝试融合边缘特征处理不规则、模糊的边界问题以达到最先进的分割性能。解码器使用跳跃连接融合多尺度信息，并添加 scSE 注意力机制进行特征监督，用以抑制无效特征，增强鲁棒特征。最后，研究针对边缘分割设计损失函数，提升对耕地斑块的识别效果。

边缘梯度部分和耕地多光谱数据、指数特征、纹理特征的多特征输入部分，使用分支子网络，相对独立地提取对应的特征信息。在处理深度特征时，深度分支网络逐层将特征图输出至边缘分支子网络对应层级，为网络增加更多的底层信息；在处理边缘特征时，边缘分支子网络经过 CoT 模块的特征提取，结合门控卷积进行特征监督，并在编码器末端进行融合多分支输出。解码器每个解码模块的输入由 2 个部分组成，分别为上一模块的输出和深度分支子网络的跳跃输出，在进行多级上采样之后，将特征图恢复至输入图像的分辨率。

BECU-Net 在以 EfficientNet 为骨干网络的 U 型编解码结构中加入边缘分支子网络，增强边缘特征输出的同时为解码器部分提供丰富的特征输入，结合跳跃连接存储耕地的纹理细节信息，以提高网络的分割能力。

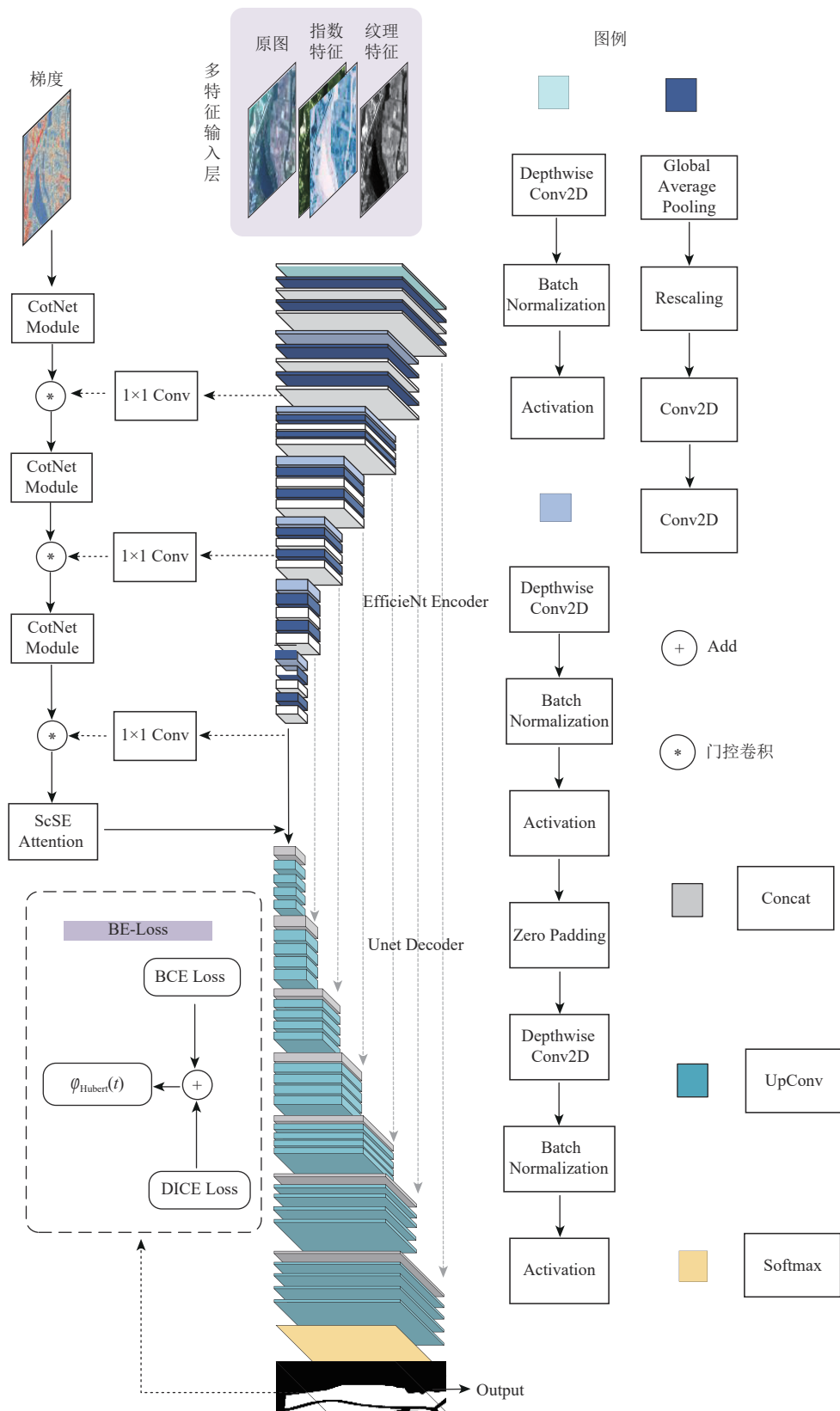


图2 本研究中改进的BECU-Net网络结构
Fig. 2 Improved BECU-Net structure in this study

2.3 边缘分支子网络

边缘检测是图像处理中最经典的问题之一，它是对象分类、分割和识别的最关键步骤，其产生

于诸如灰度、颜色和纹理等图像特征的不连续变化。在灰度图像中，边缘通常表示对象边界或物理属性（如亮度或反射率）的变化；而在多光谱图像

中, 因为光谱信息丰富使边缘更为复杂。因此, 为了更好的识别耕地遥感影像中的边界细节, 研究提出边缘分支子网络专注处理形状信息。边缘增强模块的输入分别为梯度图 ∇G 和深度分支子网络中骨干网络的编码中间结果, 实现边缘特征与深度特征双重监督, 信息互补。边缘分支中各级梯度特征图与深度特征输入级联后进行一次 1×1 卷积, 随后输入 CoT 模块融合邻近与全局的上下文信息, 并使用门控卷积提高边缘掩码的特征依赖, 最后送入 1×1 卷积统一特征图通道数并与深度特征分支的最后一级输出进行级联。在编码器的最后一层将深度特征与边缘特征融合, 最后经过门控卷积, 传入解码器的 scSE 注意力机制, 为解码网络提供更多的浅层细节信息。

2.3.1 CoT 模块

经典的自注意力机制仅是根据元素地址 query 和 key 的交互计算注意力矩阵, 而忽略了相邻 key 之间的联系。CoT 模块 (Li 等, 2023) 在二维特征图中结合相邻 key 间的上下文信息增强特征关联, 其模型将 ResNet 结构中的 3×3 卷积替换成 CoT 模块取得了较好的效果。CoT 模块运算流程如图所示。

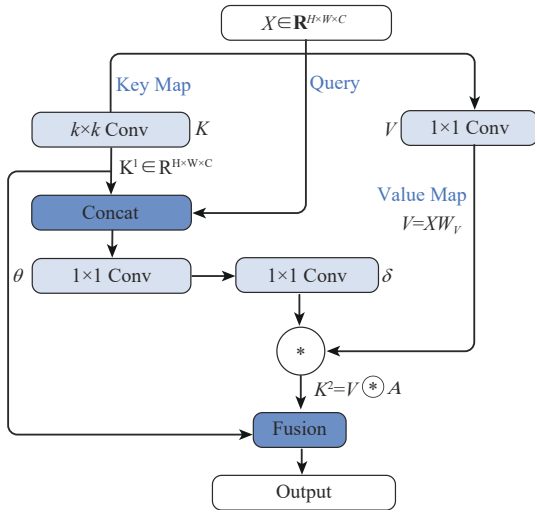


图3 CoT 计算流程

Fig. 3 CoT calculation process

模块首先对于输入特征 X , 在 K 上进行 $k \times k$ 的分组卷积, 获得具备局部上下文信息表示的 K^1 。随后使用将 K^1 和 Q 进行级联, 通过 2 个连续的卷积得到注意力矩阵 $A = [K^1, Q]W_\theta W_\delta$ 作为全局的上下文信息, 最终的输出结果为通过邻近上下文信

息与全局上下文信息的融合得到的增强特征 $K^2 = V \circledast A$, 从而充分挖掘上下文信息, 完成动态上下文表达, 增强模型对上下文信息的感知能力。

2.3.2 门控卷积

门控机制在网络中可提高边缘特征的依赖, 屏蔽深度特征对边界信息的干扰, 从而获得具有更强边缘效应的区域特征图。在 Gated-SCNN (Takikawa 等, 2019) 中, GCL 模块 (Gated Convolutional Layer) 就利用相应的门控机制帮助形状提取分支过滤信息, 实现常规流和形状流之间的特征监督。

基于此, 本文在边缘增强分支网络中设计使用了 3 组门控卷积进行边缘特征增强, 对不同训练阶段的特征进行过滤, 形成对更高级别语义信息的理解, 使边缘增强分支能以简单的结构实现高效率的图像分割。其计算公式如下:

$$\alpha_t = \sigma \left(C_{1 \times 1} (s_t \parallel r_t) \right) \quad (3)$$

$$\hat{s}_t^{(i,j)} = (s_t \circledast w_t)_{(i,j)} = \left(\left(s_{t_{(i,j)}} \odot \alpha_{t_{(i,j)}} \right) + s_{t_{(i,j)}} \right)^T w_t \quad (4)$$

式中, r_t 和 s_t 分别表示深度特征分支和边缘增强分支的中间输出, $\alpha_t \in R^{n \times w}$ 代表多分支的特征图注意力映射, $C_{1 \times 1}$ 表示 1×1 卷积, $C_{1 \times 1}$ 之后是一个 sigmoid 函数 σ ; \parallel 表示特征图级联, \odot 表示点积, w_t 表示通道权重卷积核。

2.3.3 scSE 注意力机制

scSE (Roy 等, 2018) 注意力模块通过探索空间与通道中各自的内部相关性, 捕获丰富的上下文信息。本文对多分支编码器的输出计算注意力权值, 首先对解码器的输入特征图进行全局池化和 1×1 卷积, 将其转换成通道维度的向量, 经 sigmoid 函数归一化之后得到权值掩码, 然后对特征图进行逐像素的信息校准乘法得到 cSE 通道注意力结果, 最后直接进行 $1 \times 1 \times 1$ 的卷积将特征图降维至一维, 经由 sigmoid 激活函数之后得到空间注意力特征图, 并直接与原特征图进行校准。在解码器第一层末尾对 2 种注意力进行融合处理, 得到最终的特征输入。

2.4 损失函数

损失函数在模型中起着举足轻重的作用, 并且是优化器的重要组成部分。其决定模型训练的走向。选择正确的损失函数可以帮助模型专注学习正确的特征集, 以实现最佳和更快的收敛。因此,

为了保留感兴趣的耕地边缘扩展特性,设计了一种基于边缘增强的损失函数,增强边界像素的特

征响应,即边缘增强损失函数 BE-Loss (Boundary Enhancement Loss)。

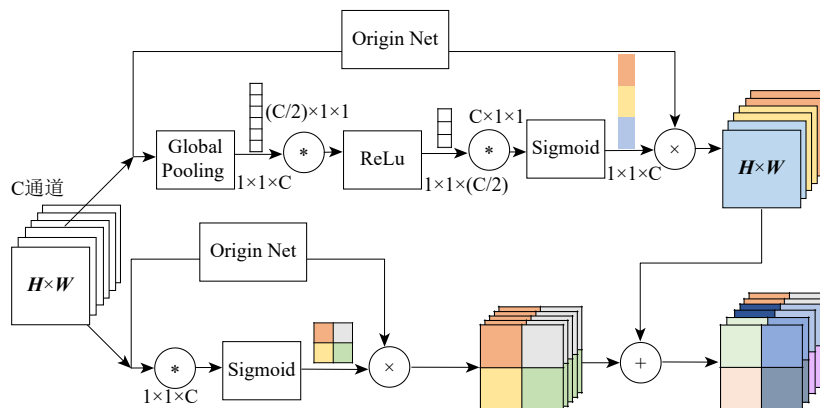


图4 scSE注意力机制

Fig. 4 scSE attention mechanism

二值交叉熵损失 BCE-Loss 通常使用 one-hot 编码计算模型预测与真实标签的相似程度。二值交叉熵损失算法原理如图 5 所示,其计算公式如下:

$$\mathcal{L}_{\text{BCE}}(X, \hat{y}) = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})) \quad (5)$$

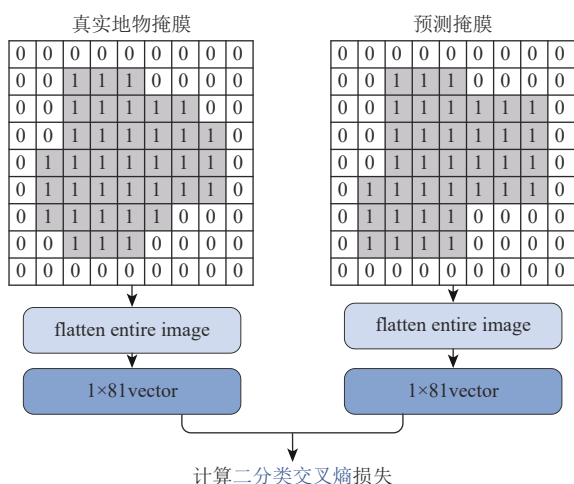


图5 BCE损失计算流程

Fig. 5 BCE loss calculation process

在本文中,对于边缘增强损失,强调对耕地边缘准确预测,研究使用 opencv 开源算法 (Zhu 等, 2020) 计算预测图像的双阈值分割结果,分别对应于耕地的内外边界,并使用骨架提取算法对边缘进行细化,最后将其附加到对应的 one-hot 编码中进行损失计算 (图 6)。

然而,自然地理的地表覆盖中,耕地所占比例相对较少,单纯使用 BCE-Loss 会导致模型训练

处理表格二进制数据,使用 Dice 系数构建的损失函数 (Li 等, 2020) 在二值分割中效果明显,适用于处理类别不平衡等问题, Dice-Loss 公式为

$$\text{Dice} = 1 - \frac{2|X \cap Y| + \text{smooth}}{|X| + |Y| + \text{smooth}} \quad (6)$$

式中, X 和 Y 分别表示真实标签和预测掩膜,为防止分母为 0, 式中加入 smooth 项。仅单独使用 Dice-Loss 时,生成的分割在高精度的同时往往表现出召回分数低的缺陷。因此,使用 BCE-Loss 和 Dice-Loss 构建联合损失函数可在增强边缘的同时避免模型训练类别不平衡问题。

此外,虽然上述联合损失函数对边缘特征已具有较好的响应,但在训练中仍然存在被非边缘像素主导损失函数的可能。故本文使用一种边缘检测算子的 Huber 正则阈值选择方法 (张善立 等, 2015), 确保模型预测的边界像素与真实标签在不相匹配时受到惩罚。正则化项 $R(\mu)$ 通过对图像中相邻像素间灰度的差值进行加权求和计算求得,具体如下:

$$R(\mu) = \sum_j R(\mu) = \sum_j \sum_{k \in S_j} \omega(k, j) \varphi(\mu_j - \mu_k) \quad (7)$$

式中, j 为图像域中所有像素的索引, S_j 表示二维空间第 j 个像素点的邻域。权重 $\omega(k, j)$ 表示像素 k 和像素 j 相互关系的正常数,为正值并且是对称的,即 $\omega(k, j) \geq 0$ 且 $\omega(k, j) = \omega(j, k)$ 。 $\omega(k, j)$ 通常被设定为邻域 S_j 内像素 k 和 j 之间距离的倒数。 φ 表示一个凸的正势函数并且满足 $\varphi(0) = 0$, 在本文中,该函数为 Huber 正则项,公式如式 (6)。

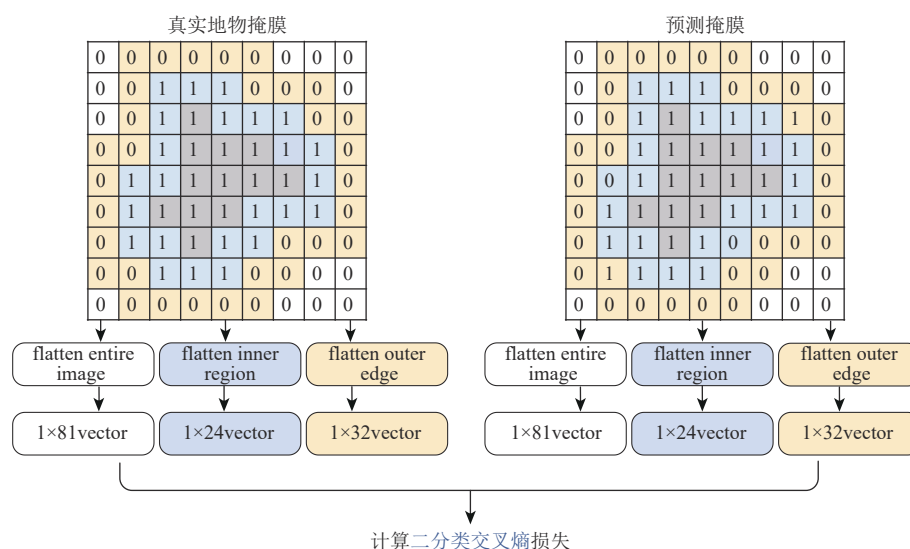


图6 BE损失计算流程

Fig. 6 BE-loss calculation process

$$\varphi_{\text{Huber}}(t) = \begin{cases} t^2/2, & |t| \leq \delta \\ \delta|t| - \delta^2/2, & |t| > \delta \end{cases} \quad (8)$$

式中， δ 代表边缘提取阈值。Huber正则化的主要优点是它能根据图像当前像素点与其相邻像素点间的灰度差值来调节对不连续区域的惩罚，这一作用由其阈值 δ 实现。

本文所设计的联合损失函数可以有效的去除噪声，令边缘增强模块更专注处理边界信息，产生可更清晰的预测对象边界，并显著提高对细小对象的识别能力。

3 数据获得与处理

3.1 实验数据及实验平台

3.1.1 实验数据

GID (Tong等, 2020) 是第一个典型的陆地覆盖分类带标签数据集，遥感图像分辨率高达4 m。为各项高空间分辨率遥感图像的陆地覆盖分类研究提供高质量的可靠数据集。该数据集包含来自中国60多个不同城市的150张高质量的高分二号影像，高分二号影像自2014年启动以来，因其兼备高空间分辨率和宽视场优势，被广泛用于土地覆盖调查、环境监测、作物估算、建设规划等重要工作中。GID大规模分类数据集广泛分布于5万多公里的地理区域。且具有多时态特点，包含了不同季节同一区域地物的差异性光谱特征。数据集参考了中国土地使用分类标准 (GB/T21010—

2017)，确定了一个层次类别系统。数据集中标注了5个主要类别：建筑、农田、森林、草地和水，分别标注了红色、绿色、青色、黄色和蓝色。不属于上述5类的区域和杂波区域被标记为黑色背景。在本文中使用相关“农田”影像和标签作为实验原始数据经预处理后进行训练与测试。

3.1.2 实验平台

本文的实验平台采用torch1.7与cuda 11.1的训练框架，硬件使用英特尔 i7-9300 CPU 和英伟达 3080ti GPU。

训练过程中的激活函数选用 ReLu，最大学习率设置为0.1，Batchsize 设置为16，最大迭代次数为100。衰减权重设为2。优化器选用 Adam，并将 Gamma 值设为2。

3.2 数据预处理

GIDGF-2数据集共含150张7200×6800像素的高分辨率大尺度数据。表1为数据集序列号。本文从中选取不同季节、不同地理位置的原始数据，最终获得8736组测试数据，2912组验证数据，2912组测试数据。对标签文件中的多类目标进行二值化，整理得到仅含耕地、非耕地2个类别的标签集，二值化结果中，白色代表耕地，黑色代表非耕地。

四波段影像与多特征计算后的特征图融合，作为网络输入，特征包括MSAVI指数、GBSI指数和由灰度共生矩阵生成的纹理特征图 (图7)。

表1 GIDGF-2数据集中选取的实验数据

Table 1 Selected experimental data from the GID-GF2 dataset

季节	训练集	验证集	测试集
春	0001517689,0001537716,0001537637	0001642620	0001591676
夏	0001015646,0001660727,0001680851	0001757429	0001757317
秋	0001064454,0001246644,0001876775	0001886305	0001723991
冬	0000647768,0000564691,0001367840	0001465929	0000564539

注:各串数字为在GID数据集中关于原始影像的序列号中纯数字部分。

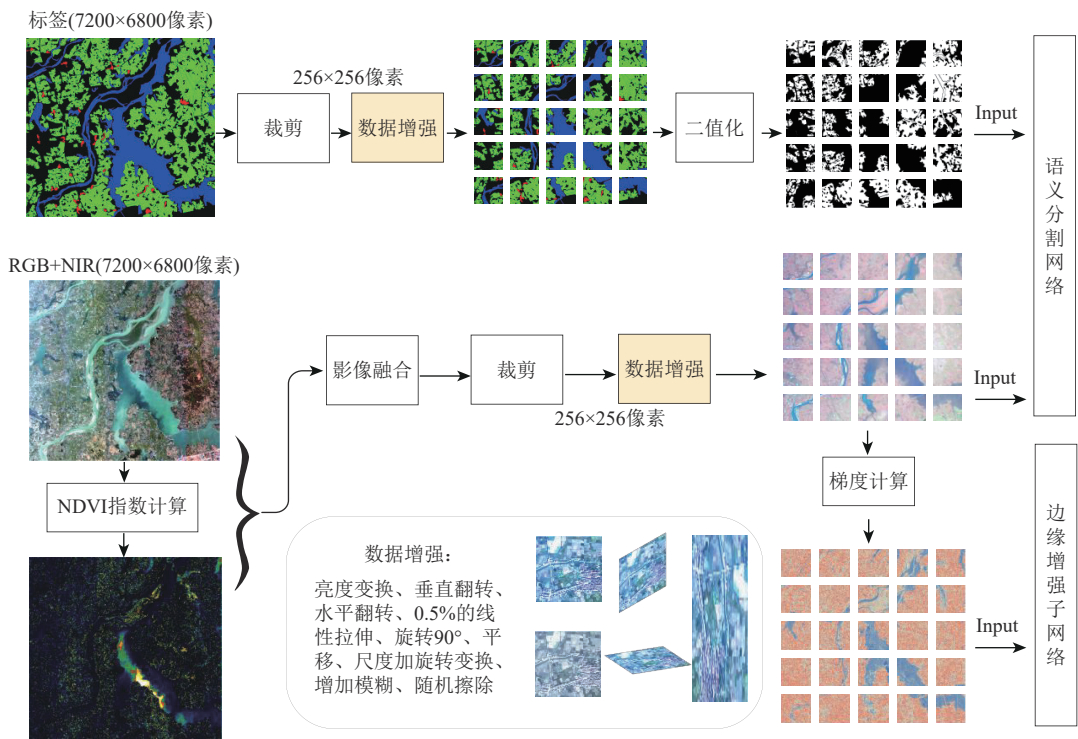


图7 遥感影像预处理

Fig. 7 Remote sensing image preprocessing

3.3 评价指标

本文实验所用主要精度指标价为准确率 (Accuracy)、精确率 (Precision) 和召回率 (Recall), 计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

式中, TP、TN、FP、FN 分别表示被正确分为正例、正确分为负例、错误分为正例、错误分为负例的个数。

F1 分数 (F1-Score) 指标同时兼顾分类模型的精确率和召回率, 为精确率和召回率的调和平均数, 因此本文使用 F1 分数作为参考指标。其计

算公式为

$$F_1 - Score = 2 \frac{Precision \cdot Recall}{Precision + Recall} \tag{12}$$

此外, 增加图像分割评价的常用指标交并比

IoU:

$$IoU = \frac{TP}{TP + FP + FN} \tag{13}$$

4 结果和分析

4.1 语义分割多模型对比分析

利用处理后的GID数据, 分别对6种主流模型进行测试, 并在测试完成后从实验结果中选取具有耕地代表性类型的5组提取效果图用于结果分析 (图8)。可见不同模型的预测结构与标签对比有如下特征:

- (1) 第1组 (第1行) 对照实验中部分区域存

在覆膜耕地, 由于在 GID 数据集中覆膜耕地并未被归类为耕地, 故本文关于实验结果的评价分析遵循此套标准。其他各组模型均能不同程度上破碎化地反映了该区域覆膜耕地的存在和分布情形, 而 BECU-Net 识别结果整块地划出了覆膜耕地所在范围。第 2 组对照实验中, PANet、LInkNet、PSPNet 存在明显的细节缺失, 对于耕地道路间隔的提取欠缺, 容易出现遗漏、边缘闭合。而 BECU-Net 对于耕地分界线分割完整, 道路明确。第 3 组对照实验中, 对于休耕的自然耕地和裸土情景, BECU-Net 分类准确, 而 PANet、LInkNet、PSPNet 出现误检和强噪声现象, Unet++ 的分割不完整, DeepLabV3+ 在细节边缘分割不全。第 4 组对照实验中, PSPNet 出现误检现象, 而 BECU-Net 在针对耕地边界的道路、居民区识别中效果明显优于其他网络。第 5 组对照实验中, PANet、LInkNet、PSPNet 均出现误检现象, 用红圈标出 BECU-Net 改善了其他网络的共同错判区域的识别情况。第 6 组对照实

验中, 耕地占比较小, 识别难度较大, 随着网络复杂度的提升, 识别效果越来越好, 其中 DeeplabV3+ 识别效果较好, 但提取结果边缘粗糙, 且出现噪声点; 而 BECU-Net 整体识别效果边缘清晰, 无误检。

(2) 从总体上, PANet、LinkNet、PSPNet 提取的边界出现碎屑斑块, 边界线粗糙, 且较其他模型误差较大; Unet++ 提取的耕地边缘出现部分区域误提现象, 对于耕地间隙 (尤其是道路、绿化带) 提取不完整; DeepLabV3+ 则在部分地区的细节部分提取缺失, 锯齿状边界现象更明显。BECU-Net 模型所提取的耕地的边缘分明, 与耕地真实边缘形状接近, 空洞和孤岛噪声少, 提取的大型地块没有缺失且边角锐利, 提取的小型地块轮廓清晰, 形变小, 在各类间隙和复杂边缘处, 对于 GID 数据集的提取效果较其他 5 种模型具有明显改善, 用于边缘提取时效果显著, 有效抑制了耕地斑块的锯齿和空洞现象。

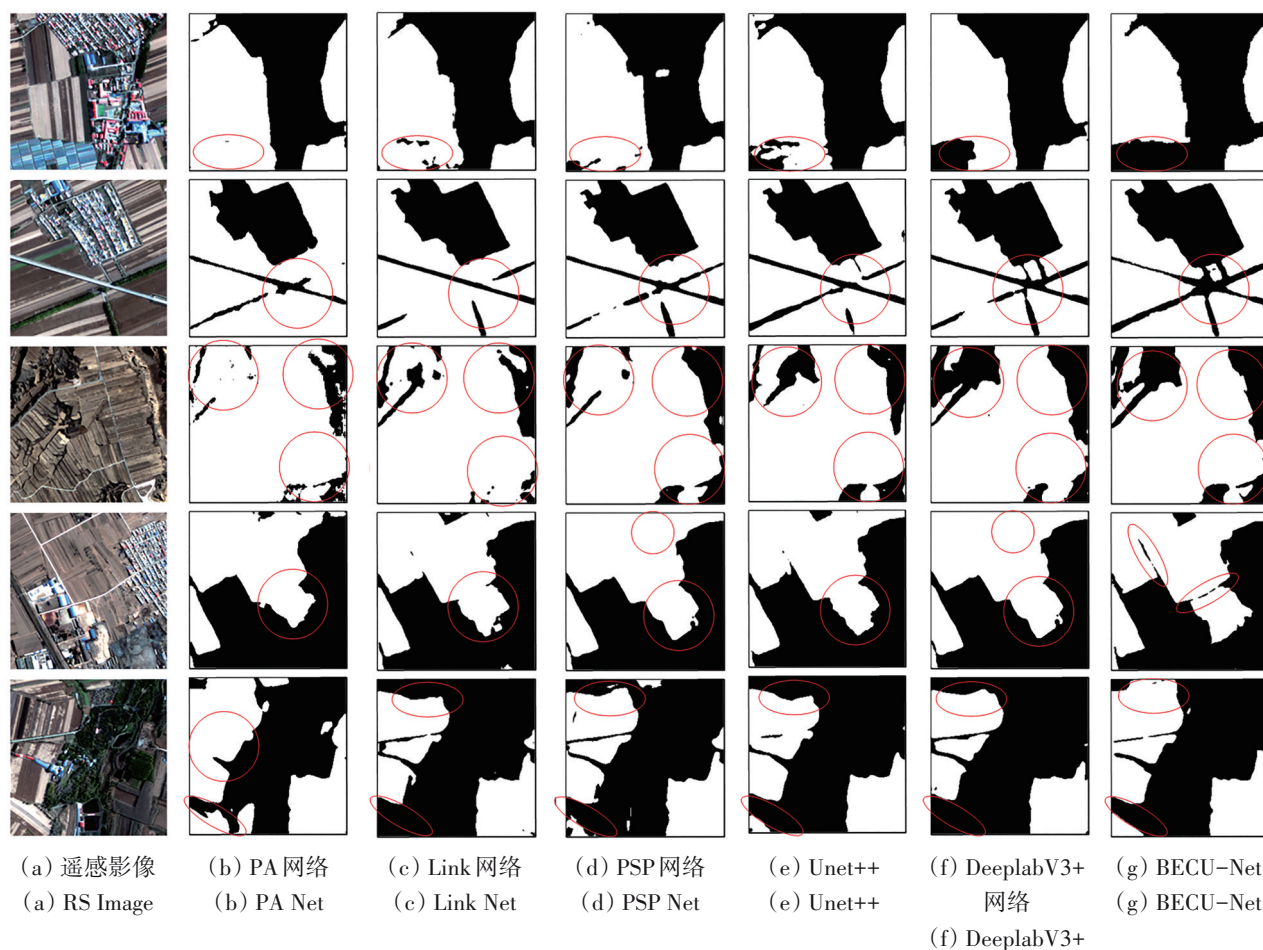


图8 不同网络的提取效果对比

Fig. 8 Comparison of extraction effects of different networks

为了定量评价本文方法的有效性, 利用 PANet、LInkNet、PSPNet、DeepLabV3+、Unet++、BECU-Net 模型在 GID 数据集上进行测试, 统计了不同模型的 IOU、召回率、F1-score 等指标, 结果如表 2 所示。可见, BECU-Net 网络精度达 93.42%, 较 PANet、LInkNet、PSPNet、Unet++、DeepLabV3+ 分别提升 9.58%、7.68%、6.49%、3.11% 和 2.72%; 对于耕地的提取准确率为 94.34%, 相较于 PANet、LInkNet、PSPNet、Unet++、DeepLabV3+ 分别提升 15.22%、14.22%、8.57%、3.79% 和 2.24%; 召回率为 96.98%, 相对 PANet、LInkNet、PSPNet、Unet++、DeepLabV3+ 分别提升 16.77%、10.87%、6.75%、4.07%、3.14%, 较 PANet、LInkNet、PSPNet、Unet++、DeepLabV3+ 分别提升 13.91%、8.80%、10.41%、3.08% 和 2.56%; F1-score 为 95.17%, 较 PANet、LInkNet、PSPNet、Unet++、DeepLabV3+ 分别提升 8.55%、5.40%、6.37%、3.42% 和 1.77%。

表 2 不同网络模型的实验精度结果对比

Table 2 Comparison of experimental accuracy results of different network models

网络模型	精度	Precision	Recall	IoU	F1-score
PANet	0.8384	0.7912	0.8021	0.7736	0.8724
LinkNet	0.8574	0.8012	0.8611	0.8247	0.9039
PSPNet	0.8693	0.8577	0.9023	0.8086	0.8942
Unet++	0.9031	0.9055	0.9291	0.8819	0.9237
DeeplabV3+	0.9070	0.9210	0.9384	0.8871	0.9402
BECU-Net	0.9342	0.9434	0.9698	0.9127	0.9579

各指标表明, BECU-Net 较其他网络有明显提升, 对精度和召回率等指标均有一定的效果。与经典网络相比, 本文构造的网络模型结合耕地遥感指数和纹理特征, 并设计特征子网络充分利用耕地边缘信息, 有效提高网络识别耕地的能力。

训练时的损失变化情况反映了模型的收敛速度和稳定程度等性能。图 9 和图 10 分别为不同网络模型的 F1-score 和 Loss 随批次变化的情况。其中 BECU-Net 网络曲线在训练过程中表现出更高的精度、低损失和更稳定的特点, 模型收敛效果更佳。

4.2 损失函数对比分析

对使用的几种损失函数进行测试, 其结果如图 11 所示。可见: 使用边缘增强损失函数 BE-loss

的模型较其他损失函数更为优秀, 更接近于真实地表覆盖, 分割的精细程度也更高。使用 DICE 损失函数的模型预测结果出现边缘钝化, 识别碎屑化等现象; 使用 BCE 损失的模型分割结果在难分类区域存在明显的误检; 使用联合损失 DICE+BCE 的结果细节更完善, 但存在明显漏检; 使用 BE-Loss 的结果边缘平整、准确, 无大区域错检、漏检。

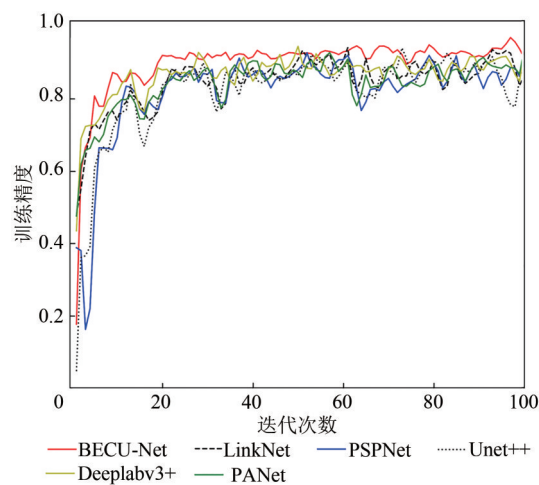


图 9 语义分割多模型训练精度(F1 分数)

Fig. 9 Model training accuracy of Semantic segmentation (F1-score)

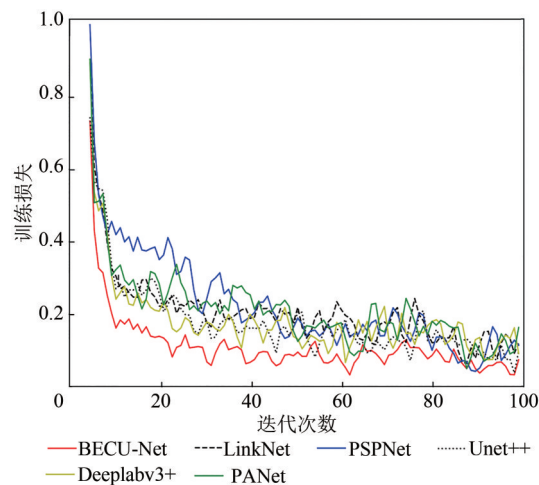


图 10 模型训练损失

Fig. 10 Model training loss

为了获取直观的损失函数对比结果, 使用各个损失函数的最优模型在测试集上进行模型评估, 各实验中的测试指标如表 3 所示。可见使用 Beloss 的模型在各项指标均高于其他模型, 在模型精度、召回率、IoU、F1-score 等指标上较联合损失分别提升 0.80%、1.92%、0.49% 和 0.56%。

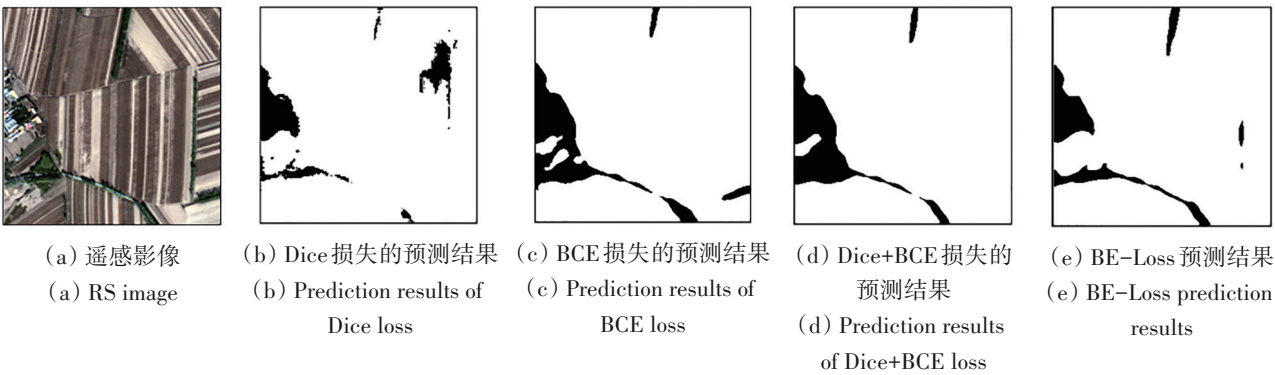


图 11 不同损失函数提取精度对比
Fig. 11 Comparison of extraction accuracy of different loss functions

表 3 不同损失函数的精度对比
Table 3 Comparison of accuracy of different loss functions

损失函数	精度	Recall	IoU	F1-score
DICE Loss	0.9126	0.9317	0.8937	0.9437
BCE Loss	0.9203	0.9484	0.903	0.9511
DICE+BCE Loss	0.9262	0.9506	0.9078	0.9523
BE-Loss	0.9342	0.9698	0.9127	0.9579

5 结 论

本文针对耕地提取结果碎片化、边界模糊等问题，构建耕地指数、纹理的多特征输入层，设计使用梯度输入的边缘分支子网络，辅以改进的损失函数，实验结果表明，提出的方法有效提高耕地信息提取精度，尤其是复杂边缘的分类准确度。在 GID 数据集上的分类结果 Precision 达 94.34%、F1-Score 达到 95.79%。相比目前现有的提取方法，BECU-Net 具有以下优势：

(1) 网络构造多样特征的输入层，包括指数特征和纹理特征，可有效反映耕地特性；(2) 边缘分支子网络专注处理形状信息，更好地识别耕地影像中的边界细节，其边缘特征与 Efficient 编码器的深度特征信息互补，二者级联可充分利用浅层细节信息；(3) 改进的含正则项的联合型损失函数 BE-Loss，解决训练样本类别不平衡、非边缘像素主导损失函数的问题。

本研究不足之处在于：GID 数据集标签为像素级分类结果，标签覆膜准确性仍有欠缺。后续研究希望建立对立的非监督学习提取方法，进一步解决神经网络对标签的依赖，实现真正的全自动耕地信息提取。

参考文献 (References)

Bao Y T, Liu W, Gao O Y, Lin Z K and Hu Q. 2021. E-Unet++: a semantic segmentation method for remote sensing images// 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference. Chongqing: IEEE: 1858-1862 [DOI: 10.1109/IMCEC51613.2021.9482266]

Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation//Proceedings of the 15th European Conference on Computer vision. Munich: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2_49]

Chen Z X, Wang L M, Wu W B, Jiang Z W and Li H. 2016. Monitoring plastic-mulched farmland by Landsat-8 OLI imagery using spectral and textural features. Remote Sensing, 8(4): 353 [DOI: 10.3390/rs8040353]

Du G M, Gai Z X and Wang H Y. 2021. Theoretical explanation and research framework of cultivated-land fragmentation in China. Journal of Earth Sciences and Environment, 43(6): 997-1008 (杜国明, 盖兆雪, 王洪彦. 2021. 中国耕地细碎化的理论解析与研究框架. 地球科学与环境学报, 43(6): 997-1008) [DOI: 10.19814/j.jese.2021.07029]

Fan D L, Su X Y, Weng B, Wang T S and Yang F Y. 2021. Research progress on remote sensing classification methods for farmland vegetation. AgriEngineering, 3(4): 971-989 [DOI: 10.3390/agriengineering3040061]

Gan F P, Wang R S, Wang Y J and Fu Z W. 1999. The classification method based on remote sensing techniques for land use and cover. Remote Sensing for Land and Resources, 11(4): 40-45 (甘甫平, 王润生, 王永江, 付正文. 1999. 基于遥感技术的土地利用与土地覆盖的分类方法. 国土资源遥感, 11(4): 40-45) [DOI: 10.6046/gtzyyg.1999.04.08]

Gong P and Howarth P J. 1990. The use of structural information for improving land-cover classification accuracies at the rural-urban fringe. Photogrammetric Engineering and Remote Sensing, 56(1): 67-73

He C, Li S L, Xiong D H, Fang P Z and Liao M S. 2020. Remote sensing

- image semantic segmentation based on edge information guidance. *Remote Sensing*, 12(9): 1501 [DOI: 10.3390/rs12091501]
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc.: 1097-1105
- Lee R Y, Chang K C, Ou D Y and Hsu C H. 2020. Evaluation of crop mapping on fragmented and complex slope farmlands through random forest and object-oriented analysis using unmanned aerial vehicles. *Geocarto International*, 35(12): 1293-1310 [DOI: 10.1080/10106049.2018.1559886]
- Li X Y, Sun X F, Meng Y X, Liang J J, Wu F and Li J W. 2020. Dice loss for data-imbalanced NLP tasks. *arXiv preprint arXiv*: 1911.02855 [DOI: 10.48550/arXiv.1911.02855]
- Li Y H, Yao T, Pan Y W and Mei T. 2023. Contextual transformer networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2): 1489-1500 [DOI: 10.1109/TPAMI.2022.3164083]
- Liu G Y, Song X and Lv J. 2015. Farmland parcels extraction from high-resolution remote sensing images based on the two-stage image classification//Proceedings Volume 9812, MIPPR 2015: Automatic Target Recognition and Navigation. Enshi: SPIE: 239-244 [DOI: 10.1117/12.2209212]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Ma M N. 1987. Basic theories and methods of remote sensing visual interpretation. *Remote Sensing Information*, (3): 26-29 (马霭乃). 1987. 遥感目视解译的基本理论与方法. *遥感信息*, (3): 26-29
- Mather P M. 1986. Review of: "Introductory Digital Image Processing: A Remote Sensing Perspective" By J. R. JENSEN: (Englewood Cliffs, New Jersey: Prentice Hall, 1986) [Pp. 368.] Price £53-45. *International Journal of Remote Sensing*, 7(12): 1836-1838 [DOI: 10.1080/01431168608948975]
- Pan Y Z, Hu T G, Zhu X F, Zhang J S and Wang X D. 2012. Mapping cropland distributions using a hard and soft classification model. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11): 4301-4312 [DOI: 10.1109/TGRS.2012.2193403]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Roy A G, Navab N and Wachinger C. 2018. Concurrent spatial and channel 'squeeze and excitation' in fully convolutional networks//21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada: Springer: 421-429 [DOI: 10.1007/978-3-030-00928-1_48]
- Takikawa T, Acuna D, Jampani V and Fidler S. 2019. Gated-SCNN: gated shape CNNs for semantic segmentation. *arXiv*: 1907.05740 [DOI: 10.48550/arXiv.1907.05740]
- Tan M X and Le Q V. 2020. EfficientNet: rethinking model scaling for convolutional neural networks. *arXiv*: 1905.11946 [DOI: 10.48550/arXiv.1905.11946]
- Tong X Y, Xia G S, Lu Q K, Shen H F, Li S Y, You S C and Zhang L P. 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment*, 237: 111322 [DOI: 10.1016/j.rse.2019.111322]
- Verma A K, Garg P K and Prasad K S H. 2017. Sugarcane crop identification from LISS IV data using ISODATA, MLC, and indices based decision tree approach. *Arabian Journal of Geosciences*, 10(1): 16 [DOI: 10.1007/s12517-016-2815-x]
- Wu Y R, Wang W Z, Zhuang J X, Ma C F, Liu S H and Wu L Z. 2013. Extraction of saline land based on decision tree approach using Landsat TM DATA//2013 IEEE International Geoscience and Remote Sensing Symposium. Melbourne: IEEE: 3762-3765 [DOI: 10.1109/IGARSS.2013.6723649]
- Wu Z H, Lei S G, Bian Z F, Huang J and Zhang Y. 2019. Study of the desertification index based on the albedo-MSAVI feature space for semi-arid steppe region. *Environmental Earth Sciences*, 78(6): 232 [DOI: 10.1007/s12665-019-8111-9]
- Xia L G, Luo J C, Sun Y W and Yang H P. 2018. Deep extraction of cropland parcels from very high-resolution remotely sensed imagery//2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics). Hangzhou: IEEE: 1-5 [DOI: 10.1109/Agro-Geoinformatics.2018.8476002]
- Xu L, Ming D P, Zhou W, Bao H Q, Chen Y Y and Ling X. 2019. Farmland extraction from high spatial resolution remote sensing images based on stratified scale pre-estimation. *Remote Sensing*, 11(2): 108 [DOI: 10.3390/rs11020108]
- Xu W N, Deng X P, Guo S X, Chen J S, Sun L Y, Zheng X R, Xiong Y F, Shen Y and Wang X Q. 2020. High-resolution u-net: preserving image details for cultivated land extraction. *Sensors*, 20(15): 4064 [DOI: 10.3390/s20154064]
- Yang B and Liu Y. 2008. Technological essentials of visual interpretation of agricultural remote sensing images. *Modernizing Agriculture*, (4): 37-39 (杨博, 刘义). 2008. 农业遥感影像目视解译技术要点. *现代化农业*, (4): 37-39 [DOI: 10.3969/j.issn.1001-0254.2008.04.003]
- Yang X, Zhu D M, Yang R S, Fu Z T and Xie W B. 2020. A visible-band remote sensing index for extracting impervious surfaces. *Transactions of the Chinese Society of Agricultural Engineering*, 36(8): 127-134 (杨栩, 朱大明, 杨润书, 付志涛, 谢文斌). 2020. 一种提取不透水面的可见光波段遥感指数. *农业工程学报*, 36(8): 127-134 [DOI: 10.11975/j.issn.1002-6819.2020.08.016]
- Zhang S L, Zhang H, Hu D B, Zeng D, Bian Z Y, Lu L J, Ma J H and Huang J. 2015. Edge-detecting operator-based selection of Huber regularization threshold for low-dose computed tomography imaging. *Journal of Southern Medical University*, 35(3): 375-379 (张善立, 张华, 胡德斌, 曾栋, 边兆英, 路利军, 马建华, 黄静). 2015. 基于边缘检测算子的Huber正则化阈值选择方法在低剂量CT重建中的应用. *南方医科大学学报*, 35(3): 375-379 [DOI: 10.3969/j.issn.1673-4254.2015.03.12]
- Zheng C, Chen Y C, Shao J and Wang L G. 2022. An MRF-based multigranularity edge-preservation optimization for semantic segmentation of remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 19: 8008205 [DOI: 10.1109/LGRS.2021.3058939]

Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2020. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6): 1856-1867 [DOI: 10.1109/TMI.2019.2959609]

Zhu N Y, Liu C, Singer Z S, Danino T, Laine A F and Guo J. 2020. Segmentation with residual attention U-Net and an edge-enhancement approach preserves cell shape features. *arXiv preprint arXiv: 2001.05548* [DOI: 10.48550/arXiv.2001.05548]

Cultivated land extraction from high-resolution remote sensing images based on BECU-Net model with edge enhancement

DONG Zhangyu^{1,2,3}, LI Jinhui¹, ZHANG Jin¹, YU Jinqiu¹, AN Sen¹

1.School of Computer and Information, Hefei University of Technology, Hefei 230601, China;

2.Anhui Key Laboratory of Industrial Safety and Emergency Technology, Hefei University of Technology, Hefei 230601, China;

3.Anhui Provincial Laboratory of Intelligent Interconnection System, Hefei University of Technology, Hefei 230601, China

Abstract: Cultivated land cover, as an important technical index to reflect the dynamic changes of human activities and the utilization degree of land resources, has been widely utilized in the fields of food security assessment and land management decision making. Existing information extraction methods ignore the differential characteristics of the plots and the rich information found in edge details, which results in fragmented extraction results with fuzzy boundaries. Therefore, an improved model that couples semantic segmentation model and edge enhancement is proposed to better solve the problem of insufficient fitting of cultivated land edges and fully utilize the rich semantic features and edge information in remote sensing images. The edge loss is designed accordingly to further improve the training accuracy and model performance.

We design an edge branching self-network formed by CoT unit, gated convolution, and SCSE attention mechanism to realize the information complementarity of edge and depth features. We construct a joint edge enhancement loss function called BE-loss with constraints to enhance the attention of the model to boundary information. On this basis, we construct a cultivated land information extraction model, that is, BECU-net, by combining the EfficientNet backbone network and U-frame. In the multi-feature input layer of this model, the index and texture features of the preprocessed data are pre-extracted, the input structure is adjusted, and the feature expression ability of the network is improved.

The extraction accuracy of cultivated land is 94.13%, and the F1-score is 95.17%. Compared with PANet, the extraction accuracy increased by 15.01%, and the F1-score improved by 7.93%. Compared with DeeplabV3+ network, the extraction accuracy is enhanced by 2.03%, and the F1-score is increased by 1.15%. The edge of cultivated land extracted by BECU-Net model is clear, and it is close to the real edge shape of cultivated land. Few holes and islands are observed. The extracted large parcels are not missing, and the edges and corners are sharp. The extracted small parcels have clear outlines and small deformation. At various gaps and complex edges, the extraction effect of GID dataset is significantly improved compared with that of the five other models. The effect is significant when used for edge extraction. The sawtooth and cavity phenomena of cultivated land patches are effectively restrained as well.

(1) The input layer of network structure with multiple features, including exponential features and texture features, can effectively reflect the characteristics of cultivated land. (2) The edge branch subnetwork focuses on processing the shape information to better identify the boundary details in the cultivated land image. Its edge features complement the depth features of the Efficient encoder, and they can be cascaded to fully utilize the shallow details. (3) The improved combined loss function called BE-Loss with regular term solves the problem of unbalanced training sample categories and non-edge pixel-dominated loss function. Overall, the algorithm in this study provides a technical reference for further solving the problem of fuzzy boundaries when extracting cultivated land information. It also offers theoretical support for the accurate division of complex boundaries.

Key words: remote sensing, edge enhancement, cultivated land extraction, semantic segmentation, U-Net, high resolution remote sensing image

Supported by Anhui Natural Science Foundation Project (No.2108085MF233); Key Research and Development Program of Anhui Province (No.202004a07020030); Special Fund for Basic Scientific Research Business of Central Universities (No. JZ2021HGTB0111)