光电技术及应用

DOI：10.16818/j.issn1001－5868.2024062201

基于对抗学习与形状修正的遥感图像建筑提取

王若兰，李　辉*

（四川大学 视觉合成图形图像技术国家级重点实验室，成都 610064）

摘　要： 高分辨率遥感图像建筑物提取对城市实景三维重建具有重要意义。针对传统卷积方法分割复杂背景遥感图像中的建筑物时出现的地物互相遮挡和边界模糊导致分割精度不高问题，提出一种改进 HRNet 的遥感图像建筑物提取网络模型（SCGAN）。基于 HRNet 结构引入形状修正单元提升模型对建筑物边缘和形状的感知能力，结合对抗学习策略强化建筑物边界和几何形状等细节特征。实验结果表明，基于对抗学习和形状修正单元的 SCGAN 有效提升了分割精度，在 WHU 和 Massachusetts 数据集上的 *IoU* 分别为 90.94％和 70.89％，与流行的语义分割模型相比表现最佳。

关键词： 神经网络；语义分割；遥感图像；建筑物提取；对抗学习

中图分类号： P237　**文章编号：** 1001－5868(2024)06－1031－08

Remote Sensing Image Building Extraction Based on Adversarial Learning and Shape Correction

WANG Ruolan，LI Hui

（National Key Laboratory of Fundamental Science on Synthetic Vision，Sichuan University，Chengdu 610064，CHN）

Abstract： The extraction of buildings from high-resolution remote sensing images is of great significance in three-dimensional reconstruction of urban scenes. A remote sensing image building extraction shape correction generative adversarial network（SCGAN）with improved high-resolution network（HRNet）is proposed to address the problem of low segmentation accuracy caused by mutual occlusion and blurred boundaries of buildings in complex background remote sensing images when using traditional convolutional methods. Based on the HRNet structure，shape correction units are introduced to enhance the model's perception of building edges and shapes，and adversarial learning strategies are used to strengthen detailed features such as building boundaries and geometric shapes. The experimental results show that the SCGAN model based on adversarial learning and shape correction units effectively improves segmentation accuracy in building extraction，with *IoU* of 90.94％ and 70.89％ on the WHU and Massachusetts datasets，respectively，exhibiting the best performance compared to popular semantic segmentation models.

Key words： neural network；semantic segmentation；remote sensing images；building extraction；adversarial learning

0　引言

建筑物作为城市智能规划与建设中的基本要素之一，对城市空间布局、资源效率与社会互动等重要方面产生着影响。其中，建筑物精确提取在城市规划、环境监测和灾害评估等领域至关重要，尤其在城

收稿日期：2024－06－22.

基金项目：国家自然科学基金项目（U20A20161）.

* **通信作者：**李　辉

市三维重建中，其准确性和效率尤为关键。高分辨率遥感图像因其多光谱特性、覆盖范围广和成本效益，成为城市三维重建不可或缺的数据源。然而，尽管分辨率提高，遥感图像的解译精度并未同步提升，建筑物提取仍然面临诸多挑战[1]。例如，建筑物的多样性、光照和视角导致的遮挡与阴影，以及复杂背景中的其他地物干扰，都增加了提取的难度[2]。

随着卷积神经网络（Convolutional Neural Networks，CNN）在计算机视觉领域的突破性进展，其在遥感图像处理中的应用也日益增多，为图像分析带来了新的视角[3-4]。其中，常用的语义分割网络包括全卷积网络（Fully Convolutional Network，FCN）[5]，U-Net[6]，SegNet[7]和金字塔场景解析网络（Pyramid Scene Parsing Network，PSPNet）[8]。王恩德等[9]通过设计双通道特征提取融合的FCN网络，解决了传统CNN在特征细节提取上的不足。金澍等[10]提出改进的U-Net方法，通过引入注意力机制和空洞空间金字塔池化，显著提升了建筑物分割精度。熊彬等[11]结合Swin Transformer和特征金字塔，对PSPNet进行了改进。李紫薇等[12]提出一种改进SegNet模型，专注解决SegNet网络存在的梯度消失现象和网络参数多等问题。然而，使用CNN的研究主要集中于遥感图像中的纹理特征，而对建筑物形状建模的探讨较少。

而同时，生成对抗网络模型（GAN）[13]因其出色的表现被引入计算机视觉任务中。2016年，Luc等[14]将对抗学习思想应用于语义分割领域，采用全卷积网络和判别器网络结合的方法提升了分割精度。2022年，Li等[15]进一步提出了RegGAN模型，通过引入多尺度判别器强化边界分割的精度。潘俊杰等[16]从域自适应模型的角度出发，基于对抗学习提出了一种无监督域自适应框架，以解决域偏差问题。然而，以上工作对高分辨遥感图像中建筑物边缘的研究较少。

在多数情况下，高分辨率遥感图像中的建筑物均由紧凑的几何形构成，拥有清晰的边缘和锐利的角点，这使得它们比其他地物更易于区分。因此，本文提出了一种基于对抗学习策略的网络模型SCGAN，专注于高分辨率遥感图像中建筑物的形状与边缘建模。SCGAN通过引入形状修正单元和判别器，增强了对建筑物边缘的识别能力，并利用对抗学习中生成器与判别器的博弈，进一步提高了分割精度。通过与其他典型网络模型的对比实验，验证了本文方法的有效性。

1 研究理论与SCGAN模型

1.1 对抗学习机制

生成对抗网络是一种深度学习模型，由Goodfellow等于2014年首次提出[13]。该模型由两个关键组件构成：生成器（Generator，G）和判别器（Discriminator，D）。在训练过程中，生成器和判别器通过交替迭代进行竞争性训练。生成器的目标是生成高质量且难以区分于真实数据的高仿真数据，而判别器则致力于精确地区别真实数据和生成数据。图1所示为GAN网络的基本结构图。生成器G接收随机噪声作为输入，并生成模拟样本，这些样本随后与真实样本一起被送入判别器D。判别器D对输入样本进行分类，输出其对样本真实性的判断。该判断结果被用作损失信号，反馈至生成器G，指导其进一步训练。通过这种对抗性训练，生成器G和判别器D相互促进，逐步提升模型性能。GAN模型的优化目标由式(1)给出。

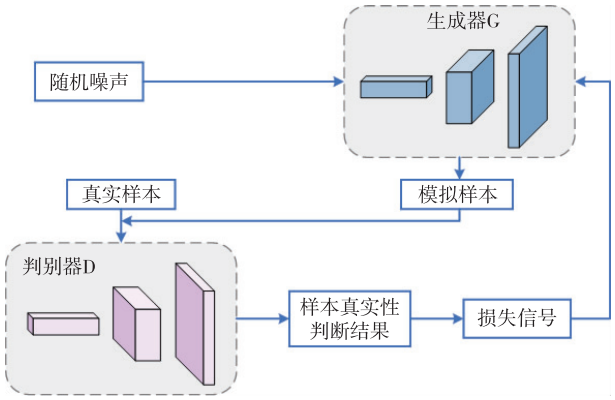


图1 GAN网络的基本结构

$$\min_G \max_D V(D,G) = E_{x\sim P_{\text{data}}}[\log D(x)] + E_{x\sim P_{z(z)}}[\log(1-D(G(z)))] \quad (1)$$

式中，E为期望，P_{data}为真实样本的分布，$P_{z(z)}$为随机噪声z的分布。当优化判别器D时，使得真实样本x的判别结果$D(x)$趋近于1，同时生成样本$G(z)$的判别结果$D(G(z))$趋近于0，即让判别器D拥有可以正确分辨真实样本和生成样本的能力，此时价值函数$V(D,G)$最大。当优化生成器G时，使得生成样本$G(z)$的判别结果$D(G(z))$趋近于1，即让生成器G拥有可以生成真实样本的能力，此时价值函数$V(D,G)$最小。在整个训练过程中，判别器D和生成器G对价值函数$V(D,G)$进行极小极

大化博弈，当 G 与 D 有足够的学习能力后，训练达到纳什均衡，给定 G，D 可以达到最优博弈解。

1.2　SCGAN 结构

所提出的 SCGAN 是一个端到端的像素级分割网络。如图 2 所示，SCGAN 的架构包括前处理、生成器和判别器三个主要部分。前处理阶段采用翻转、旋转和裁剪等方法进行数据增强，以扩充样本集并提高模型的泛化性。生成器由分割网络和形状修正单元组成，负责从遥感图像中精确提取建筑物。采用 HRNet[17] 作为分割网络的基础，该网络通过多尺度特征融合策略，有效避免了传统 CNN 在下采样过程中细节信息的损失，保持了高分辨率的特征表示，适合于对精细特征信息建模有高要求的视觉任务。形状修正单元紧接在分割网络之后，旨在引导模型学习建筑物边缘信息并进行形状正则化，从而提高模型对建筑物边缘的分割精度。生成的预测分割图像与人工标注图像一同输入至判别器，判别器专注于建筑物形状建模，通过排除背景信息，提高了对生成图像的判别能力。将判别器输出反馈至生成器，指导其进一步优化分割网络的训练过程。

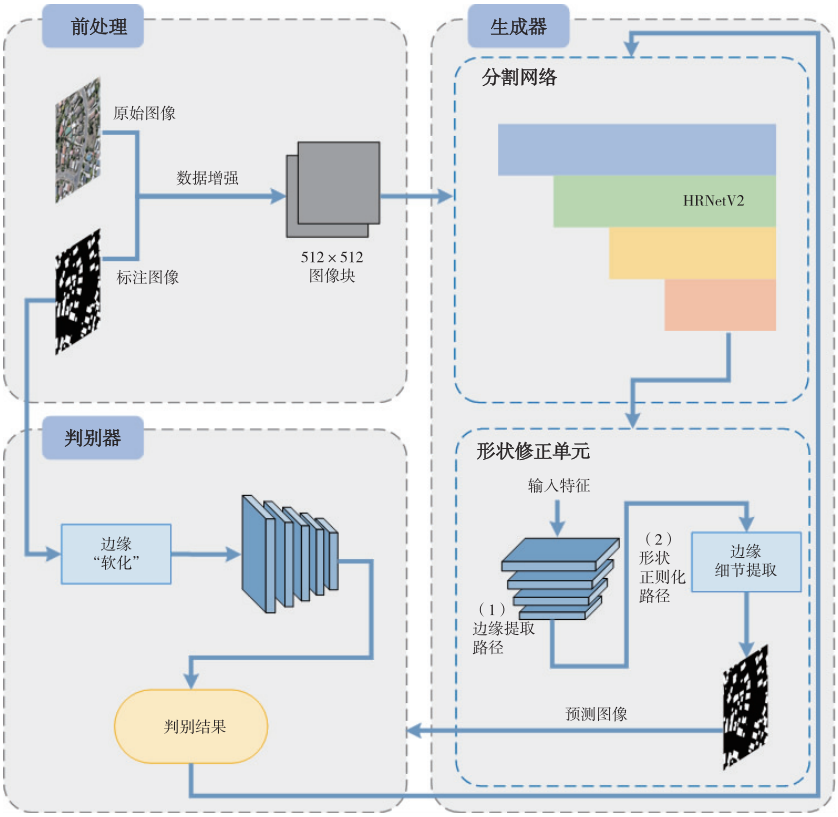


图 2　SCGAN 整体架构

损失函数用于反映模型预测结果与真实值之间的差异，直接影响模型的性能和学习效率。所提出的 SCGAN 的混合损失函数表达式为：

$$L_{\mathrm{SCGAN}}=L_{\mathrm{G}}+L_{\mathrm{D}} \tag{2}$$

式中，L_{SCGAN} 为 SCGAN 的整体损失函数，L_{G} 和 L_{D} 分别为用于优化生成器 G 和判别器 D 的损失函数。生成器使用均方差损失函数 MSE 进行训练，而判别器使用二元交叉熵损失函数 BCE 进行训练。

1.3　形状修正单元

尽管采用高分辨率网络 HRNet 作为分割网络对建筑物形状学习具有显著效果，但为了实现更精细的空间分割，适应建筑物形状特性至关重要。为此，实验在分割网络末端引入了一个形状修正单元，以更准确地拟合建筑物的形状约束。如图 3 所示，该单元包含两条关键路径：边缘提取和形状正则化。输入的特征图经过这两个路径的处理，能够输出具有更清晰边缘和更明确角点的建筑物分割预测图。

1.3.1　边缘提取路径

边缘提取路径的目标是从特征图中准确识别出预测的边缘，以此引导分割网络专注于建筑物边缘的局部细节，进而实现边缘分割的高精度。在该路径中，输入的特征图首先经历两次 3×3 卷积操作以提取关键特征，随后通过双线性上采样恢复至原始输入尺寸，确保边缘特征的完整性和准确性，完成边

缘提取过程[18]。

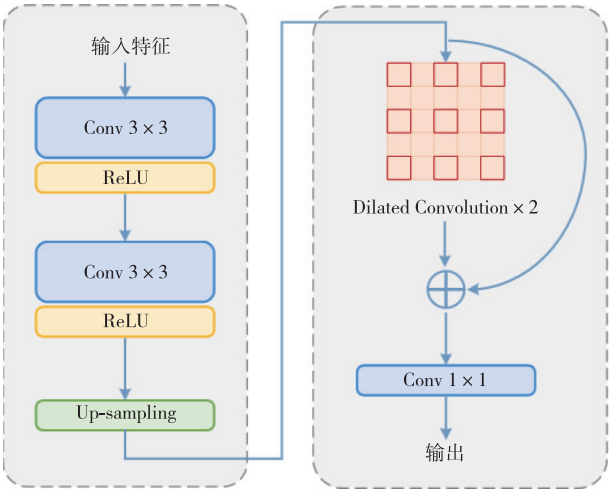


图 3 形状修正单元

1.3.2 形状正则化路径

在传统卷积中，下采样操作常导致特征图分辨率降低，丢失重要细节。为解决此问题，形状正则化路径引入膨胀卷积，通过在卷积核中插入空洞，有效扩大了感受野，使得模型能够捕获更丰富的细节信息，增强对形状特征的敏感性。采用文献[2]的方法，将两个 3×3 的膨胀卷积以残差连接的方式堆叠形成膨胀卷积组，并通过 1×1 卷积层调整特征通道数。具体操作如式(3)，设定像素 $x(r,c)$ 的卷积操作为：

$$U(r,c)=\sum_{i,j}x_{r+i,c+j}\cdot k_{i,j} \tag{3}$$

式中，$k_{i,j}$ 表示内核权重，在标准 3×3 卷积中 $i,j\in\{-1,0,1\}$，而在 3×3 膨胀卷积中 $i,j\in\{-r,0,r\}$，其中 r 为膨胀率。

1.4 判别器

在多数情况下，对抗生成网络的判别器仅在图像级别区分图像的真伪。本文提出了一种基于 CNN 的多尺度判别器，该判别器能够在像素级别区分预测图像与标注图像。如图 4 所示，该判别器通过连续的层级结构逐步提取复杂特征，以增强模型对遥感图像中建筑物边缘和形状细节的敏感性。鉴于高分辨率遥感图像中常见的地物繁杂、类间遮挡和光影干扰等问题，本文避免了直接使用遥感图像作为判别器输入，以减少背景冗余信息的干扰，确保模型专注于建筑物形状特征的学习。考虑到人工标注图像的边缘清晰度与生成器输出的分割图像之间存在差异，直接使用这些图像可能导致判别器建模失败。因此，采用下采样操作对标注图像中的建筑物边缘进行“软化”处理，减少与生成图像的对比度差异。随后，图像通过 5 组 3×3 卷积、归一化、激活和池化操作，进一步提取和精炼建筑物的形状信息。最终，通过 1×1 卷积层调整特征通道数，输出判别结果。

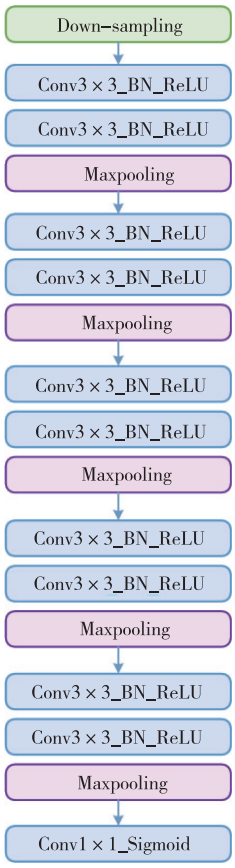


图 4 判别器

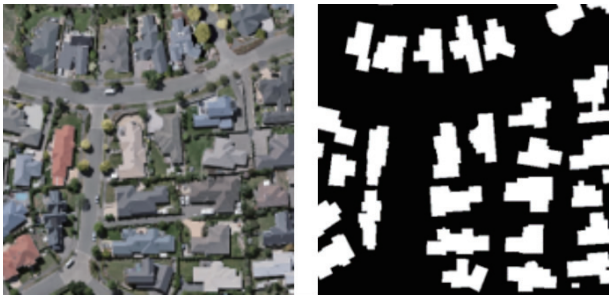
2 实验结果与分析

2.1 实验环境

本文实验平台为 Intel(R) Xeon(R) Silver 4214R CPU@2.40 GHz，GPU 使用带有 12 GB 显存的 RTX 3080 Ti，软件配置为 Ubuntu18.04 系统以及 Pytorch 深度学习框架。训练过程中使用 Adam 优化算法[19]进行优化，训练批次大小设置为 8，迭代 50 次，初始学习率设置为 0.000 1，且在训练过程中动态调整学习率，使模型更好地收敛。

2.2 数据集介绍

为确保对所提出方法 SCGAN 的有效性进行准确验证和评估，实验采用 WHU 建筑物数据集[20]进行实验验证，如图 5 所示。



（a）原始图像　　（b）标注图像

图 5　WHU 建筑物数据集

WHU 建筑物数据集是武汉大学开源的大型高分辨率建筑物检测数据集，该数据集由 22 万多栋独立建筑组成，从新西兰克赖斯特彻奇地区提取，占地 450 km^2，包含农村、城镇、文化工业区等多种区域，具有大小形态不一的建筑类型。原始影像经过下采样处理后地面分辨率为 0.3 m，并裁剪成 8 189 张 512×512 像素的图像，其中训练集包含 4 736 张、验证集包含 1 036 张、测试集包含 2 416 张。

2.3　评价指标

为更准确地量化验证所提出的方法对建筑物的分割效果，选取精确度（P）、召回率（R）、交并比（IoU）和 F_1 分数作为评价指标。其中，P 表示预测为正例中实际为正例的比例，R 表示实际为正例中被预测为正例的比例，IoU 表示标签真实值与预测值的交集和并集的比值，F_1 分数表示精确率和召回率的调和平均，全面反映了分类模型的准确性与完整性。以上计算公式分别为：

$$P=\frac{N_{TP}}{N_{TP}+N_{FP}} \tag{4}$$

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}} \tag{5}$$

$$IoU=\frac{N_{TP}}{N_{TP}+N_{FP}+N_{FN}} \tag{6}$$

$$F_1=2\times\frac{P\times R}{P+R} \tag{7}$$

式中，N_{TP} 为正确识别建筑物的像素个数，N_{FP} 为错误将背景像素识别为正确的建筑物像素的像素个数，N_{FN} 为错误将建筑物像素识别为背景像素的像素个数。

2.4　消融实验分析

为验证所提网络 SCGAN 及其各个组件的有效性，进行了消融实验。所提网络 SCGAN 选取了 HRNet 作为基准模型，添加了形状修正单元和判别器。在消融实验中，使用 WHU 建筑物训练集对网络进行训练，并通过测试集来评估分割效果，同时采用 F_1 分数和 IoU 作为评价指标来量化网络表现，实验结果如表 1 所示。

表 1　消融实验在 WHU 数据集上定量评价结果

编号	基准模型	形状修正单元	判别器	F_1 分数/%	IoU/%
1	√	×	×	93.15	88.72
2	√	√	×	94.02	90.06
3	√	×	√	94.23	90.28
4	√	√	√	94.70	90.94

表 1 所示为不同模型在遥感图像建筑物分割任务中的表现。基准模型 HRNet 的 F_1 分数和 IoU 分别为 93.15%和 88.72%。引入形状修正单元后，F_1 分数和 IoU 分别提升了 0.87 个百分点和 1.34 个百分点。进一步集成判别器，F_1 分数和 IoU 的增益分别为 1.08 个百分点和 1.56 个百分点。最终，将形状修正单元和判别器整合至基准模型，形成 SCGAN 模型，其 F_1 分数和 IoU 相较于基准模型分别显著提升了 1.55 个百分点和 2.22 个百分点。这些结果表明，形状修正单元和多尺度判别器均对提升建筑物分割精度具有显著效果。

2.5　对比实验分析

为评估所提 SCGAN 模型的分割性能，实验将其与 U-Net，SegNet，PSPNet 以及 HRNetV2 模型进行了定量对比实验。实验结果如表 2 所示，其中 SCGAN 在精确度 P、召回率 R、IoU 和 F_1 分数四项评价指标上均优于对比模型。

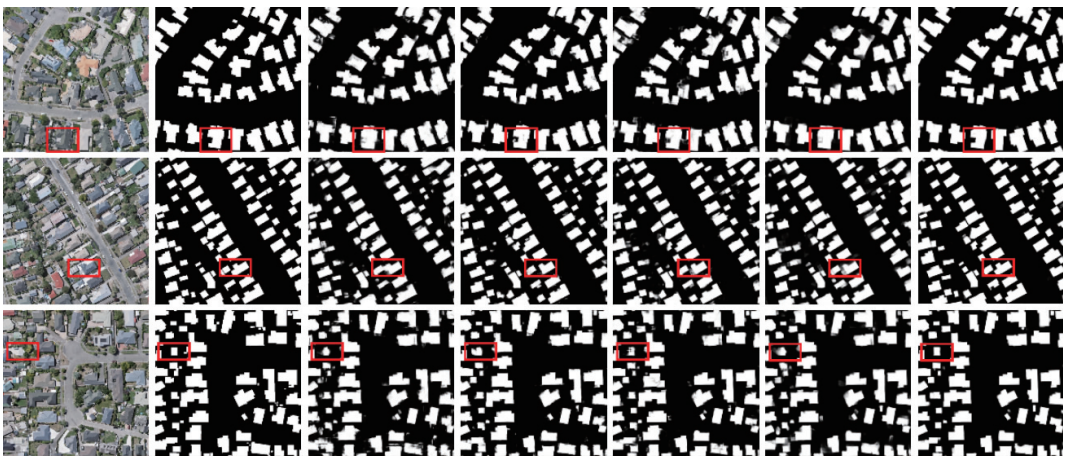
表 2　对比实验在 WHU 数据集上定量评价结果

模型	P/%	R/%	F_1 分数/%	IoU/%
U-Net	93.16	90.94	92.31	87.49
SegNet	91.57	91.68	91.14	86.64
PSPNet	93.67	92.56	93.08	88.74
HRNetV2	94.10	93.12	93.15	88.72
SCGAN	95.19	94.82	94.70	90.94

U-Net 采用编码器-解码器架构，通过跳跃连接整合多层次特征，最终分割结果的 IoU 为 87.49%。SegNet 利用其独特的编码器-解码器结构及池化索引，展现了其在精确分割和模型轻量化方面的优势，最终分割结果的 IoU 为 86.64%。PSPNet 通过金字塔池化模块捕获多尺度信息，平衡了局部细节与全局信息，最终分割结果的 IoU 为 88.74%。基准模型 HRNet，通过交叉连接多尺度特征，保持了高分辨率特征图，体现了其在细节和上下文信息准确性上的优势，最终分割结果的 IoU 为 88.72%。

SCGAN 在 WHU 数据集上的 *IoU* 达到 90.94%，为所有对比模型中最高。与基准模型 HRNet 相比，SCGAN 在精确度 P、召回率 R、F_1 分数和 *IoU* 上分别提升了 1.09 个百分点、1.7 个百分点、1.55 个百分点和 2.22 个百分点。SCGAN 通过引入对抗学习机制和形状修正单元，显著增强了对建筑物形状和边缘的感知能力，四项指标的全面提升验证了所提方法的有效性。

为了从视觉质量角度进一步验证所提方法的有效性，选取了测试集中具有代表性的 3 个局部场景，并对不同模型的预测结果进行了直观比较。这些场景中的建筑物在大小、形状、位置和光照条件上存在显著差异。如图 6 所示，黑色区域代表背景，白色区域代表建筑物，而红色矩形框突出了重点分析区域。对比结果表明，对于受遮挡和阴影影响、轮廓边界难以识别的建筑物，U-Net，SegNet，PSPNet 和 HRNetV2 的分割性能均存在局限。相比之下，所提出的 SCGAN 模型，得益于对抗学习机制和形状修正单元的引入，在小建筑物与大建筑物相连区域的分割上表现出更高的准确性。此外，SCGAN 在提取被树木等其他地物遮挡的建筑物边界方面也展现了优越性。对比实验结果表明，SCGAN 在最大化建筑物形状规则性和轮廓清晰度方面具有优势。



（a）原始图像 （b）标注图像 （c）U-Net （d）SegNet （e）PSPNet （f）HRNetV2 （g）SCGAN

图 6 对比实验在 WHU 数据集上可视化评价结果

2.6 泛化能力分析

为验证所提 SCGAN 模型的泛化能力，选择了较低分辨率的 Massachusetts 建筑物数据集[21]进行验证，确保模型不受限于单一数据集的特定结构和特征。Massachusetts 数据集的空间分辨率为 1 m，与 WHU 数据集相比，Massachusetts 数据集的质量和分辨率都较低。

WHU 与 Massachusetts 数据集的实验结果对比如表 3 所示，相较于 WHU 数据集，SCGAN 在 Massachusetts 数据集上的识别性能有所下降，精确度 P、召回率 R 与 *IoU* 分别下降 8.76%，15.11% 和 20.05%。

表 3 WHU 与 Massachusetts 数据集实验结果对比

数据集	P/%	R/%	*IoU*/%
WHU	95.19	94.82	90.94
Massachusetts	86.43	79.71	70.89

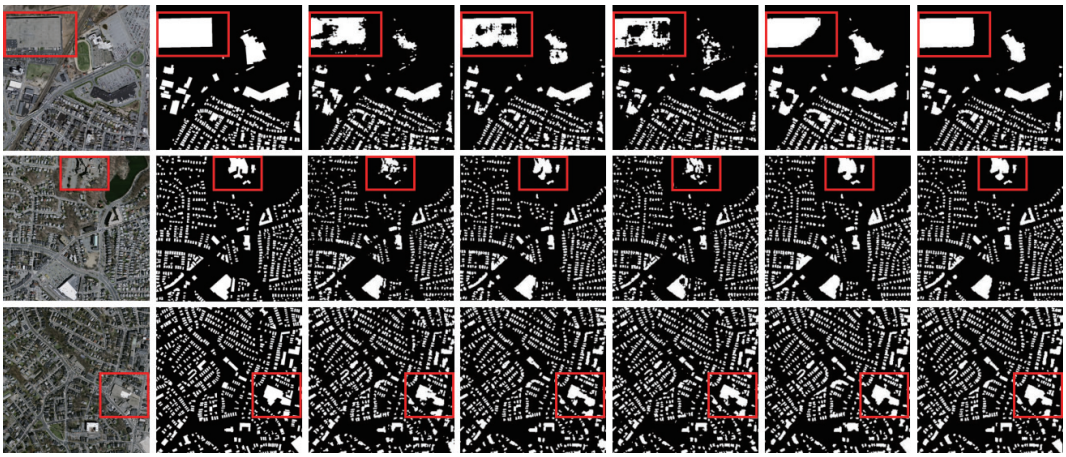
在 Massachusetts 数据集上，将所提方法 SCGAN 与 U-Net，SegNet，PSPNet 和 HRNetV2 进行定量对比实验，其建筑物分割结果如表 4 所示。从表 4 可以看出，所提出的 SCGAN 模型在精确度 P、召回率 R、F_1 分数与 *IoU* 四项指标上均优于 U-Net，SegNet，PSPNet 和 HRNetV2。与基准模型 HRNetV2 相比，SCGAN 的精确度 P 提升了 1.66 个百分点，召回率 R 提升了 1.11 个百分点，F_1 分数提升了 1.77 个百分点，*IoU* 值提升了 2.58 个百分点。实验结果表明，SCGAN 在高分辨率和低分辨率数据集上都能提升建筑物提取精度。

表 4 对比实验在 Massachusetts 数据集上定量评价结果

模型	P/%	R/%	F_1 分数/%	*IoU*/%
U-Net	79.62	72.43	75.81	61.14
SegNet	83.95	76.94	80.25	67.09
PSPNet	83.31	74.27	79.69	66.54
HRNetV2	84.77	78.60	81.12	68.31
SCGAN	86.43	79.71	82.89	70.89

图 7 为不同网络模型在 Massachusetts 测试集上部分影像的分割结果对比图，重点分析区域以红

色矩形框表示。由图 7 可以看出，在建筑物边缘及与周围地物相接的复杂区域，传统网络模型的分割精度不足。相比之下，所提出的 SCGAN 模型在这些区域实现了显著的分割效果改善。



（a）原始图像　（b）标注图像　（c）U-Net　（d）SegNet　（e）PSPNet　（f）HRNetV2　（g）SCGAN

图 7　对比实验在 Massachusetts 数据集上可视化评价结果

3　结论

为解决传统卷积神经网络在建筑物提取中的边界分割不精确问题，提出了一种依据深度学习的遥感图像建筑物提取算法。针对建筑物的边缘和形状特征，基于 HRNet 模型引入形状修正单元及对抗学习策略，形状修正单元能够更好地拟合建筑物的形状约束，对抗学习策略通过生成器和判别器的互相博弈，排除背景冗余信息，强化细节特征，不断优化分割结果。

实验结果表明，相较于基准模型 HRNet 以及其他对比模型，所提出的 SCGAN 方法对遥感图像中建筑物的形状及边界信息具有更高的敏感度，在建筑物的边缘易混淆处、建筑物之间以及建筑物被其他地物遮挡的相连处具有较好的分割性能。尽管如此，对于形状不规则、具有非直线边界的建筑物，SCGAN 的分割精度仍需进一步提高。未来的研究将集中于提高非规则建筑物分割的精确度。

参考文献：

[1] Li J, Huang X, Tu L, et al. A review of building detection from very high resolution optical remote sensing images[J]. GIScience & Remote Sensing, 2022, 59(1): 1199-1225.

[2] Ding L, Tang H, Liu Y, et al. Adversarial shape learning for building extraction in VHR remote sensing images[J]. IEEE Transactions on Image Processing, 2021, 31: 678-690.

[3] 王志敏，王加胜，王丽蒙. 基于注意力机制的 U 型网络遥感影像分割[J]. 计算机仿真，2023，40(3)：232-235，240.
Wang Zhimin, Wang Jiasheng, Wang Limeng. Remote sensing image segmentation based on attention mechanism U-net[J]. Computer Simulation, 2023, 40(3): 232-235, 240.

[4] 陈天华，郑司群，林宇骁. 基于改进深度神经网络的遥感影像语义分割[J]. 计算机仿真，2021，38(12)：27-32.
Chen Tianhua, Zheng Siqun, Lin Yuxiao. Remote sensing image semantic segmentation based on improved deep neural networks[J]. Computer Simulation, 2021, 38(12): 27-32.

[5] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.

[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, 2015: 234-241.

[7] 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.

[8] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6230-6239.

[9] 王恩德，齐　凯，李学鹏，等. 基于神经网络的遥感图像语义分割方法[J]. 光学学报，2019，39(12)：93-104.
Wang Ende, Qi Kai, Li Xuepeng, et al. Semantic segmentation methods for remote sensing images based on neural networks[J]. Acta Optica Sinica, 2019, 39(12): 93-104.

[10] 金　澍，关　沫，边玉婵，等. 基于改进 U-Net 的遥感影像建筑物提取方法[J]. 激光与光电子学进展，2023，60(4)：59-65.
Jin Shu, Guan Mo, Bian Yuchan, et al. Building extraction method for remote sensing images based on improved U-Net[J]. Progress in Laser and Optoelectronics, 2023, 60(4): 59-

65.

[11] 熊 彬，张双德. 基于改进 PSPNet 的卫星遥感图像建筑物语义分割算法[J]. 遥感信息，2023，38(4)：73-79.

Xiong Bin，Zhang Shuangde. Building semantic segmentation algorithm for satellite remote sensing images based on improved PSPNet[J]. Remote Sensing Information，2023，38(4)：73-79.

[12] 李紫薇，英昌盛，于晓鹏，等. 基于改进 SegNet 模型的遥感图像建筑物分割[J]. 吉林大学学报(理学版)，2022，60(2)：409-416.

Li Ziwei，Ying Changsheng，Yu Xiaopeng，et al. Building segmentation of remote sensing images based on improved SegNet model[J]. Journal of Jilin University：Science Edition，2022，60(2)：409-416.

[13] Goodfellow I J，Pouget-Abadie J，Mirza M，et al. Generative adversarial nets [J]. Advances in Neural Information Processing System，2014，3(11)：1-9.

[14] Luc P，Couprie C，Chintala S，et al. Semantic segmentation using adversarial networks[J]. arXiv preprint arXiv：1611.08408，2016.

[15] Li Q，Zorzi S，Shi Y，et al. RegGAN：An end-to-end network for building footprint generation with boundary regularization[J]. Remote Sensing，2022，14(8)：1835.

[16] 潘俊杰，慎 利，鄢 薪，等. 一种基于对抗学习的高分辨率遥感影像语义分割无监督域自适应方法[J/OL]. 自然资源遥感：1-10[2024-05-06].

Pan Junjie，Shen Li，Yan Xin，et al. An unsupervised domain adaptation method for semantic segmentation of high-resolution remote sensing images based on adversarial learning[J/OL]. Remote Sensing of Natural Resources：1-10[2024-05-06].

[17] Wang J，Sun K，Cheng T，et al. Deep high-resolution representation learning for visual recognition [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence，2020，43(10)：3349-3364.

[18] 杨潇宇，汪西莉. 结合多尺度注意力和边缘监督的遥感图像建筑物分割模型[J]. 激光与光电子学进展，2022，59(22)：335-344.

Yang Xiaoyu，Wang Xili. Building segmentation model for remote sensing images combining multi-scale attention and edge supervision[J]. Progress in Laser and Optoelectronics，2022，59(22)：335-344.

[19] Kingma D P，Ba J. Adam：A method for stochastic optimization[J]. arXiv preprint arXiv：1412.6980，2014.

[20] Ji S，Wei S，Lu M. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set [J]. IEEE Transactions on Geoscience and Remote Sensing，2018，57(1)：574-586.

[21] Mnih V. Machine learning for aerial image labeling [D]. Toronto：University of Toronto，2013.

作者简介：

王若兰(1997—)，女，重庆市人，硕士研究生，主要研究领域为遥感图像处理；

李 辉(1970—)，男，四川省人，硕士生导师，教授，主要研究领域为智能计算和虚拟现实。