

# 生成对抗学习式半监督遥感影像生成网络地图方法

伍江江<sup>1</sup>, 宋洁琼<sup>2\*</sup>, 田纪龙<sup>1</sup>, 陈浩<sup>1</sup>, 沙志超<sup>1</sup>, 李军<sup>1</sup>, 彭双<sup>1</sup>, 杜春<sup>1</sup>

(1. 国防科技大学 电子科学学院, 湖南长沙 410073; 2. 军事科学院 国防科技创新研究院, 北京 100071)

**摘要:**针对现有全监督学习获取精确配对样本耗费资源问题,同时兼顾网络地图生成质量,提出了一种新颖的基于生成对抗网络的半监督网络地图生成模型,旨在利用少量精确配对的数据和大量非配对数据,实现智能化遥感影像直接生成网络地图。此外,设计了一种基于变换一致性正则化和样本增强一致性的半监督学习策略,克服了非精确配对数据带来的不一致性问题,同时能获得更好的模型泛化性能。对不同地图数据集进行了充分的对比实验,模型生成的网络地图在定量指标和视觉质量上优于比较方法,验证了半监督网络地图生成方法的有效性和快速性。

**关键词:**网络地图生成;生成对抗网络;遥感影像;半监督学习;一致性正则化

中图分类号: P283.8 文献标志码: A 文章编号: 1001-2486(2025)03-128-13



论文  
拓展

## Online map generation method from remote sensing images via semi-supervised adversarial learning

WU Jiangjiang<sup>1</sup>, SONG Jieqiong<sup>2\*</sup>, TIAN Jilong<sup>1</sup>, CHEN Hao<sup>1</sup>, SHA Zhichao<sup>1</sup>, LI Jun<sup>1</sup>, PENG Shuang<sup>1</sup>, DU Chun<sup>1</sup>

(1. College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China;

2. National Innovation Institute of Defense Technology, Academy of Military Science, Beijing 100071, China)

**Abstract:** To address the resource consumption issue of obtaining precise paired samples in existing fully supervised learning, while also considering the quality of network map generation, a novel semi-supervised online map generation model based on generative adversarial networks was proposed, which aimed to realize the direct generation of intelligent remote sensing images into network maps by using only a few precisely matched data and a large amount of unpaired data. In addition, a semi-supervised learning strategy based on transformation consistency regularization and sample enhanced consistency was designed, which overcame the inconsistency problem caused by imprecise paired data and derives better generalization performance of the model. Adequate comparison experiments were conducted on different map datasets. The generated online maps outperform the competing methods on the quantitative metrics and visual quality, which validate the effectiveness and speed of semi-supervised network map generation methods.

**Keywords:** online map generation; generative adversarial networks; remote sensing images; semi-supervised learning; consistency regularization

随着信息技术和互联网技术的飞速发展,为地图发展提供了新的机遇,网络地图也作为新一代地图产品脱颖而出,成为人们日常生活中不可缺少的一部分。其在公共交通、自动驾驶、国土测绘等领域中发挥着重要作用。

在数字制图环境下,遥感影像因其获取便捷、成本低、更新快、适应强等特点,成为制作网络地图的重要数据来源并得到了广泛应用<sup>[1]</sup>。遥感

影像可反映出道路、水域、建筑等地物信息的分布情况,从中可以提取到绘制网络地图所需的地理信息,为各类地图的编制提供了可靠的地理信息源<sup>[2]</sup>。

传统的遥感影像到网络地图的制作过程简要包含“影像地理要素提取—矢量数据更新—地图制图综合—瓦片生成与组织”等流程<sup>[3-4]</sup>。主要依赖于人工获取矢量数据和专家经验,难度较大

收稿日期: 2024-05-11

基金项目: 国家自然科学基金资助项目(62106276, 42471403)

第一作者: 伍江江(1982—),男,四川广安人,副教授,博士, E-mail: wujiangjiang@nudt.edu.cn

\*通信作者: 宋洁琼(1993—),女,山东东营人,助理研究员,博士, E-mail: sjq@nudt.edu.cn

引用格式: 伍江江, 宋洁琼, 田纪龙, 等. 生成对抗学习式半监督遥感影像生成网络地图方法[J]. 国防科技大学学报, 2025, 47(3): 128-140.

Citation: WU J J, SONG J Q, TIAN J L, et al. Online map generation method from remote sensing images via semi-supervised adversarial learning[J]. Journal of National University of Defense Technology, 2025, 47(3): 128-140.

且过程较为复杂。再者,传统的遥感影像到网络地图的制作过程较长,相对遥感影像的更新速度而言,网络地图服务发布的地图信息可能与遥感影像不匹配,导致时效性不足,在紧急情况下较难做到迅速反应和更新。

近年来,深度学习技术得到了迅猛发展,其在图像处理<sup>[5-6]</sup>、目标识别与检测<sup>[7-8]</sup>、图像生成<sup>[9-10]</sup>等方向取得了较大的进展。已有许多专家和研究者将基于深度学习的智能化方法和模型应用在遥感影像制作网络地图过程的部分步骤中,例如应用深度学习模型对影像地理要素信息进行识别和语义分割等<sup>[11-12]</sup>,实现矢量要素的提取。虽然将深度学习技术引入传统的网络地图制作流程中,可以产生精度较高的网络地图,但仅实现了部分步骤自动化。目前仍然存在依赖高质量人工标注的像素级标签、计算和数据冗余、制作耗时较长等问题,在遇到诸如洪涝、地震、泥石流等突发自然灾害的情况下难以保证为救援人员提供地图信息的时效性。

另一类智能化方法是通过深度学习域映射技术端到端地实现遥感影像转化为网络地图。旨在使用生成模型<sup>[13-15]</sup>,学习从源图像域(遥感影像域)到目标图像域(网络地图域)的直接表示。较为流行的生成模型是2014年由古德费洛(Goodfellow)提出的生成对抗网络(generative adversarial network, GAN)<sup>[13]</sup>。基于GAN的图像域映射技术及其应用得到了迅速发展,具有代表性的框架有pix2pix<sup>[10]</sup>,其借鉴条件生成对抗网络<sup>[16]</sup>使用配对的样本学习源图像域到目标图像域的联合分布。但其要求训练数据集中源域图像与目标域图像必须成对。为打破这一约束,CycleGAN<sup>[17]</sup>、DualGAN<sup>[18]</sup>等无监督模型采用了循环一致性思想,使模型可以利用非配对的数据进行训练。然而,缺乏配对的样本作为条件监督信息使得这些无监督方法远不如全监督方法健壮。以上这些图像域映射模型为遥感影像端到端的生成网络地图提供了基础和技术支持。但是,目前基于GAN的通用域映射模型在网络地图生成的应用中存在一些局限性。

第一,基于GAN的图像域映射方法根据不同网络结构的设计,使用配对(全监督模型)或非配对(无监督模型)数据进行训练。然而,由于遥感影像和网络地图的更新周期不同步、人工绘制网络地图存在主观偏差等因素,导致利用地图服务如谷歌地图、百度地图、天地图等获取的遥感影像和对应的网络地图存在大量非精确配对样本,如

图1所示。例如,网络地图中相同地物颜色不一致、元素缺失、地物形状轮廓与影像不匹配等。大量非精确配对的样本会损害全监督模型的泛化性能。而无监督模型不需要配对的样本,但由于缺乏相应的监督信息使得其生成的网络地图精度较低。

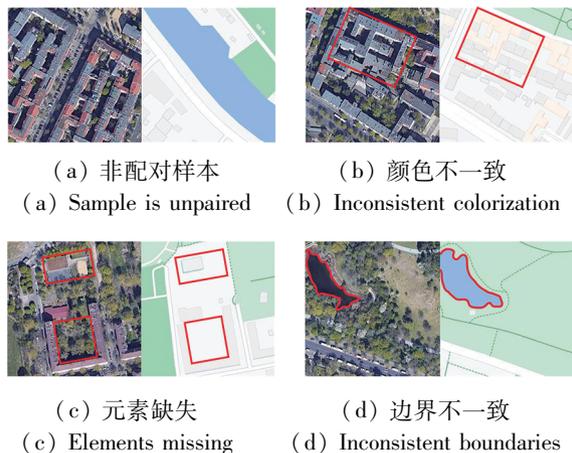


图1 非配对样本和非精确配对样本示例

Fig. 1 Example diagram of unpaired samples and imprecisely paired samples

第二,通用的基于深度学习域映射技术的图像生成模型仅采用单一的损失函数来改进输出网络地图与真实网络地图的差距<sup>[19-20]</sup>,并未针对遥感影像存在的纹理不清晰、相邻像素梯度较小等问题做出改进,从而导致某些生成的网络地图地理元素失真、局部细节模糊等问题。

因此研究拟面向同一地域不精确匹配的遥感影像与网络地图数据,采用生成对抗学习式半监督方法对网络地图生成方法展开研究。

## 1 相关工作

### 1.1 生成对抗网络

生成对抗网络<sup>[13]</sup>以其强大的建模能力、可生成以假乱真的图像,而受到国内外研究人员的极大关注。其基本思想源自双人博弈论,由一个生成器(generator)和一个判别器(discriminator)构成,通过对抗学习的方式进行训练,目的是估测数据样本的潜在分布并生成新的数据样本,这使得GAN成为近年来深度学习在复杂分布数据上进行无监督学习最具前景的方法之一,在各个领域得到迅速的发展,如图像编辑<sup>[21-22]</sup>、超分辨率<sup>[23-24]</sup>、风格迁移<sup>[25-26]</sup>、图像去噪<sup>[27-28]</sup>等。

与其他经典的生成模型如卷积神经网络(convolutional neural network, CNN)<sup>[29]</sup>、变分自编码器(variational autoencoder, VAE)<sup>[14]</sup>、流模型

(flow-based model)<sup>[30]</sup>等模型相比, GAN 具有多方面的优势: ①不需要设计遵循任何种类的因式分解模型, 生成对抗式网络框架能训练任何一种生成器网络和判别器网络; ②无须利用马尔可夫链反复采样, 无须在学习过程中进行推断, 回避了近似计算棘手的概率难题; ③相比其他所有模型, GAN 可以产生更加清晰, 真实的样本。

## 1.2 域映射

近期, 在计算机视觉领域基于深度生成模型的域映射技术得到了广泛的关注, 图像域映射技术采用自动编码(auto-encoder)的形式, 可以以无监督学习的方式实现从源图像域到目标图像域端到端的转换, 研究者以全监督或无监督的方式对图像映射分别开展了研究。

对于全监督的域映射方法, 模型需要配对的数据进行训练。pix2pix<sup>[10]</sup>是一种典型的图像域映射方法, 其借鉴条件生成网络<sup>[16]</sup>并使用配对的样本学习图像到图像的共同分布。该方法已被用于许多图像转换案例, 例如, 自动驾驶数据集 cityscapes<sup>[31]</sup>的标签到街景, 边缘素描图到照片<sup>[32]</sup>, 遥感影像或航拍图像到地图<sup>[33]</sup>等。Wang 等<sup>[34]</sup>提出的 pix2pixH 用于高分辨率的图像到图像转换, 较 pix2pix 能够生成分辨率较高的图像。文献 [35] 提出了 PLDT (pixel-level domain transfer) 的方法, 其添加了一个判别器 Dpair 以监督图像转换任务。虽然生成了高质量的结果, 但由于有监督的网络设计, 模型必须在配对数据上进行训练。

然而, 现实中的大多数任务都缺乏可用的配对样本。为了克服这一限制, 国外研究人员基于循环一致性理论提出了 CycleGAN<sup>[17]</sup>、DualGAN<sup>[18]</sup>、DiscoGAN<sup>[36]</sup>等无监督技术, 旨在图像经过两个相反的生成器转换两次后, 应该正确重建与输入图像相同的分布。文献 [37] 将 VAE<sup>[14]</sup>与基于循环一致性损失的 GAN 相结合, 假设两个编码器共享相同的隐藏空间, 为无监督图像域映射提供了新的思路。

以上这些经典的图像域映射模型为遥感影像自动化生成网络地图提供了基础和技术支持。

## 1.3 基于域映射的网络地图生成技术

基于深度生成模型的域映射方法在图像相关领域表现出了良好的效果。但遥感图像与一般图像相比, 其包含的内容远比普通图像多, 具有波段多、范围广、变形较大等特点<sup>[38]</sup>。国内外众多研究者尝试研究了针对遥感影像的专门域映射方

法, 并取得了丰硕的成果。

2019 年, Gangul 等<sup>[33]</sup>提出了全监督地图生成模型 GeoGAN, 使用条件生成网络<sup>[16]</sup>将遥感影像转换为对应的地图, 并在 GAN 损失中添加重构损失和内容损失学习地图的风格。同年, Kang 等<sup>[39]</sup>通过配置 pix2pix 和 CycleGAN 模型, 将现有的地图样式标准(如谷歌地图、OpenStreetMap)中的样式元素转移到非风格化的地理信息系统(geographic information system, GIS)矢量数据中。Song 等<sup>[40]</sup>基于域映射模型 CycleGAN 提出了遥感影像到网络地图快速转换的无监督模型 MapGen-GAN, 并设计了残差 U-Net 网络作为生成器, 以及几何/循环一致性损失函数约束模型不需要配对样本进行训练。随后, Song 等<sup>[41]</sup>提出了地图迁移模型, 通过结合深度对抗迁移训练方案和空间注意力机制来生成多个区域的网络地图。为进一步提升网络地图生成的精度, Chen 等<sup>[42]</sup>提出了一个半监督地图生成模型 SMAPGAN<sup>[42]</sup>, 并针对对象拓扑关系设计了损失函数和地图质量评价指标, 同时 Ding 等<sup>[43]</sup>提出了 Improved SMAPGAN 进一步优化了该模型, 但是该模型没有考虑地图的语义信息。Fu 等<sup>[44]</sup>设计了一种判别式模块和生成式模块组合的地图生成模型, 将风格和语义特征紧密结合在一起并取得了较好的效果。同时, 还提出了一种新的多层级地图生成框架(multilevel maps, MLM), 可以自动从航空图像中生成具有多个缩放级别一致的网络地图。以上基于生成模型的方法均不能较好地处理弱匹配数据导致的模型泛化性能衰减的问题。

## 2 变换一致性正则化

一致性正则化(consistency regularization, CR)在分类任务中表现出先进的性能。当输入样本受到扰动时, 一致性正则化强制模型保持对无监督样本的预测。出于 CR 的应用, 本文采用变换一致性正则化(transformation consistency regularization, TCR)的正则项, 以确保模型对图像样本几何变换的预测与模型对该图像重建的几何变换一致。

具体地, 给定  $K$  对配对的遥感影像和网络地图样本的集合  $D_s = \{(x_i, y_i) : i \in (1, 2, \dots, K)\}$ , 其中,  $x_i$  采样于遥感影像样本分布  $X$ ,  $y_i$  采样于网络地图样本分布  $Y$ 。从相同的分布中采样一组非配对的遥感影像和网络地图样本的集合  $D_u = \{(x_i, y_i) : i \in (1, 2, \dots, \delta K)\}$ , 其中  $\delta$  是半监督训练中每批使用的非配对样本与配对样本的比例。目标是利用无监督非配对数据来了解更多关于样

本分布的内在结构,从而正则化遥感影像到网络地图的映射网络 $f_{\theta}(\cdot)$ 。根据平滑假设<sup>[45]</sup>,在遥感影像到网络地图的转换中,对于一系列的几何变换集合 $T(\cdot)$ ,如果模型对第 $i$ 个非配对遥感影像 $x_i \in D_U$ 预测一个重建的网络地图 $y'_i$ ,那么该模型对其经过几何变换 $T(x_i)$ 的预测应为 $T(y'_i)$ ,即:

$$T_m(f_{\theta}(x_i)) \approx f_{\theta}(T_m(x_i)) \quad (1)$$

其中, $T_m$ 是一系列的几何变换集合中的操作,如旋转、翻转、横向纵向平移、裁剪等,一致性正则化和样本增强一致性策略,以无监督学习的方式使模型可以更好地利用庞大的非配对数据学习多样化样本分布,通过一致性损失函数约束生成器生成更准确的网络地图。同时,能够有效防止模型在少量的精确配对数据上过拟合,从而提高模型的鲁棒性和泛化性。

### 3 基于生成对抗网络的半监督遥感影像生成网络地图模型

#### 3.1 模型整体流程

基于GAN的半监督遥感影像生成网络地图模型包含一个改进的基于U型结构的生成器和一个马尔可夫判别器。生成器负责将遥感影像生

成网络地图,判别器经过训练负责分辨生成的网络地图是由生成器生成的还是真实的样本,两者通过对抗学习不断各自优化。在半监督训练过程中,包含两个训练阶段,如图2所示,首先是预训练阶段对模型进行预热,仅使用一小部分精确配对的样本有监督地训练生成器和判别器,以初始化两者的权重。然后是网络精调阶段,使用大量的易获取的非配对样本,通过变换一致性正则化和样本增强一致性策略,以无监督学习的方式使模型可以更好地利用庞大的非配对数据学习多样化样本分布,通过一致性损失函数约束生成器生成更准确的网络地图。同时,能够有效防止模型在少量的精确配对数据上过拟合,从而提高模型的鲁棒性和泛化性。

基于GAN的遥感影像生成网络地图模型通过半监督学习策略,能够有效地处理现实中获取的大量非精确配对样本,同时提高生成网络地图的几何精度和准确性。下面对两个训练阶段分别进行详细的介绍。

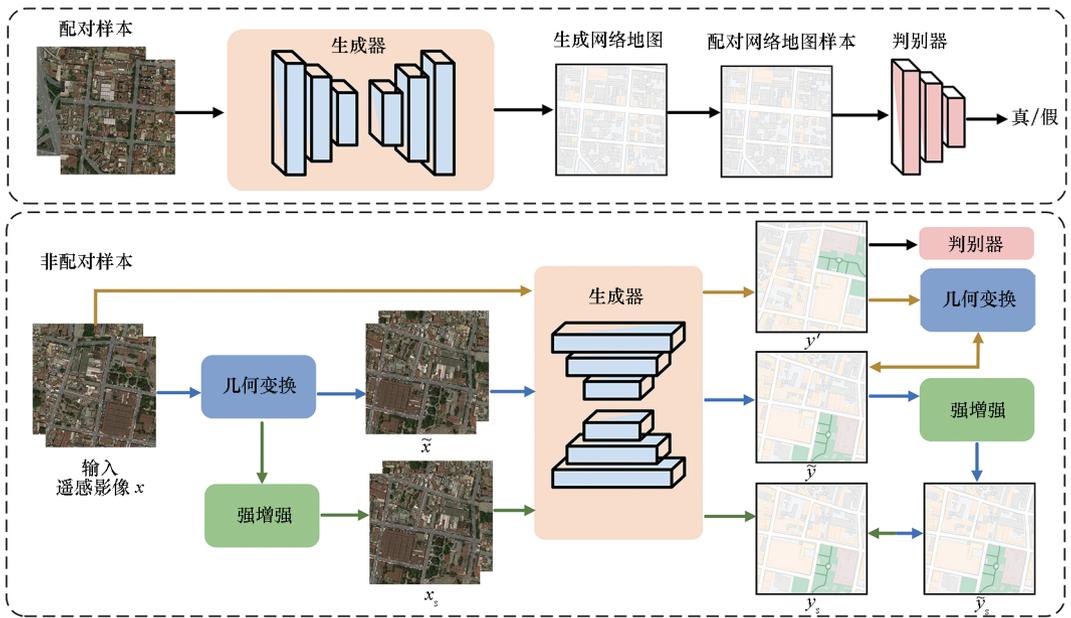


图2 基于GAN的半监督遥感影像生成网络地图模型训练流程

Fig. 2 Training flowchart of GAN-based semi-supervised remote sensing image generation network map model

#### 3.2 预训练阶段

在对基于GAN的半监督遥感影像生成网络地图模型的预训练阶段中,使用精心筛选出的精确配对遥感影像和网络地图样本对改进的U型结构生成器和马尔可夫判别器进行预热并初始化两者的权重。将遥感影像域 $X$ 中的训练样本表示为 $\{x_1, x_2, \dots, x_k\}$ ,对应的网络地图域 $Y$ 中的样本表示为 $\{y_1, y_2, \dots, y_k\}$ ,生成器的目标是学习遥

感影像域 $X$ 到网络地图域 $Y$ 的映射 $G: X \rightarrow Y$ ,而判别器试图区分生成的网络地图 $y'_i$ 是由生成器生成的还是真实的网络地图。生成器和判别器对抗式地进行训练,因此有目标函数为

$$L_{GAN}^{pre}(G, D) = E_{x \sim P(x), y \sim P(y)} [\log D(x, y)] + E_{x \sim P(x)} [\log(1 - D(x, G(x)))] \quad (2)$$

其中, $x \sim P(x)$ 和 $y \sim P(y)$ 分别为遥感影像和网络地图的数据分布, $G(\cdot)$ 表示生成器输出, $D(\cdot)$ 表示判别器输出。生成器试图最小化目标

函数,而判别器的目的是最大化目标函数,即:

$$G^* = \arg \min_G \max_D L_{GAN}^{pre}(G, D) \quad (3)$$

除欺骗判别器外,生成器的任务是产生接近真实网络地图的地图图像。因此,为促进输出  $G(x)$  更接近真实地图  $y$ ,引入一个重构损失函数,即:

$$L_{Recon}(G) = E_{x \sim P(x), y \sim P(y)} [\|y - G(x)\|_1] \quad (4)$$

其中,  $\|\cdot\|_1$  为  $L1$  距离,以减少生成网络地图模糊程度。因此,在预训练阶段,整体的目标函数为:

$$L_{pre} = \arg \min_G \max_D L_{GAN}^{pre}(G, D) + \lambda_{L1} L_{Recon}(G) \quad (5)$$

其中,  $\lambda_{L1}$  为超参数。

### 3.3 网络精调阶段

在对所有精确配对样本的预训练结束后,将非配对样本输入预训练好的生成器中进行网络精调训练阶段。通过变换一致性正则化和样本增强一致性策略使模型以无监督的方式学习多样化非配对样本的分布。

第一步,在预训练的生成器上应用变换一致性正则化策略。首先,使用几何变换函数  $T_m$  对遥感影像进行翻转、顺/逆时针旋转等操作,生成变换后的遥感影像  $\tilde{x} = T_m(x)$ ,将遥感影像输入预训练好的生成器中,生成对应的变换过的网络地图  $\tilde{y} = G(T_m(x))$ 。然后,将未经几何变换的遥感影像  $x$  输入生成器中生成网络地图  $y'$  后使用几何变换函数,对生成的网络地图进行相同的变换操作,生成变换的网络地图  $\tilde{y}' = T_m(G(x))$ ,通过变换一致性损失函数约束生成器对遥感影像几何变换的预测与对相应网络地图重建的几何变换一致,因此,变换一致性损失函数表示为:

$$L_{TCR}(G) = \frac{1}{\partial K} \sum_{i=1}^{\partial K} \|T_m(G(x)) - G(T_m(x))\|_1 \quad (6)$$

其中,  $\partial$  是半监督训练中每批次使用的非配对样本与配对样本的比例,  $K$  是数据集样本的总数,  $T_m$  是几何变换函数。变换一致性损失函数利用非配对数据来约束生成器在不同形式的几何变换上重建几何精度更高的网络地图。最后,将生成器生成的网络地图  $y'$  输入预训练的判别器中进行对抗式的训练,有对抗损失:

$$L_{GAN}^{fine}(G, D) = E_{y \sim P(y)} [\log D(y)] + E_{x \sim P(x)} [\log(1 - D(G(x)))] \quad (7)$$

第二步,在变换一致性正则化的基础上,对生成器使用样本增强一致性策略。如图 2 所示,首

先,对遥感影像进行几何变换,然后做一次图像增强处理,生成几何变化的遥感影像  $x_s$ 。强增强处理指的是能够使遥感影像产生较大变化的增强,由 2~3 种增强方法进行随机混合。每次使用时,随机选择不同的增强方法(强度随机)组合成一种强增强方法。数据增强方法旨在使模型学习到更深层次的语义信息。将使用强增强操作后的遥感影像输入相同的生成器中,生成对应的网络地图  $y_s = G(Aug(T_m(x)))$ ,  $Aug(\cdot)$  表示强增强操作;然后,将第一步中生成的变换的网络地图  $\tilde{y} = G(T_m(x))$  做一次同样的强增强操作  $Aug(\cdot)$ ,生成增强后的网络地图  $\tilde{y}_s = Aug(G(T_m(x)))$ ;使用骰子损失  $L_{dice}(y_s, \tilde{y}_s)$  保持网络地图强增强后的一致性,即:

$$L_{dice} = 1 - \frac{2 \cdot y_s \cdot softmax(\tilde{y}_s)}{y_s + softmax(\tilde{y}_s)} \quad (8)$$

其中,  $softmax(\cdot)$  是常见的归一化函数。

因此,在对基于 GAN 的半监督遥感影像生成网络地图模型精调训练阶段的损失函数为:

$$L_{fine} = \lambda_{TCR} L_{TCR}(G) + L_{GAN}^{fine}(G, D) + \lambda_{dice} L_{dice} \quad (9)$$

其中,  $\lambda_{TCR}$  和  $\lambda_{dice}$  为超参数。

### 3.4 生成器和判别器结构

U 型生成器的结构如图 3 所示,结合了跳跃连接、残差块、编-解码器结构和注意机制等元素,包含编码器、扩张桥和解码器三部分。编码器使用 ResNet34 与注意力模块,注意力机制可以关注重点区域,并帮助模型为输入的部分分配不同的权重,提取更多的关键信息;扩张桥使用了几个带有跳跃连接的扩展型卷积层,扩展的卷积层有利于扩大接受野,跳过连接有助于捕获高级和低级信息;解码器使用转置卷积层来恢复特征映射的分辨率。

判别器将网络地图作为输入,从图像中提取特征,预测其是否为真实网络地图的概率,概率值越接近 1,表明生成的网络地图越接近于真实值。为了更好地辨别输入网络地图的部分细节,采用马尔可夫判别器 PatchGAN<sup>[10]</sup> 对网络地图进行 patch 级(70×70)的区分。马尔可夫判别器由卷积层构成,输出为一个  $n \times n$  的矩阵,最后取矩阵的均值作为真/假的输出。矩阵中的每一个输出,代表着原图中一个感受野,对应了原图的一个 patch,这种 patch 级判别器结构具有更少的参数,可以以完全卷积的方式处理任何大小的网络地图。

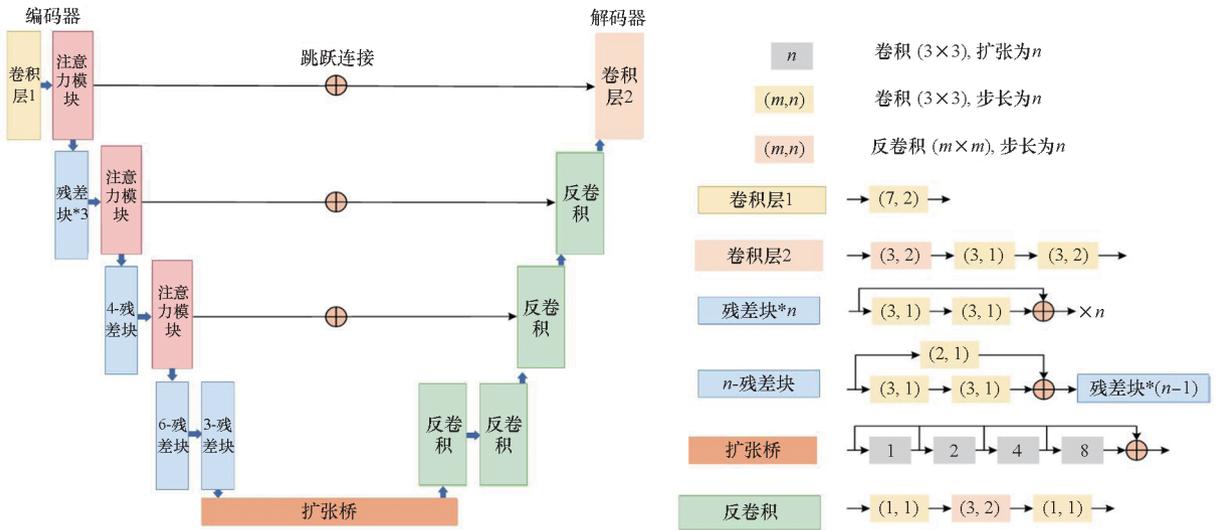


图 3 改进的 U 型生成器网络结构

Fig. 3 Structure diagram of the modified U-shape generator network

### 4 实验与分析

为了验证本文提出的基于生成对抗网络 GAN 的半监督遥感影像生成网络地图模型的有效性和可用性,本节使用三组不同地域的地图数据集和三种图像质量评价方法对提出的模型进行验证。同时,通过与几种目前通用的图像域映射方法和地图自动化生成模型进行定量、定性的对比实验,结果证明了本文提出的半监督遥感影像生成网络地图模型优于其他方法,可以生成精度较高、视觉质量较好的网络地图。

#### 4.1 数据集和评估标准

##### 4.1.1 数据集

通过大量实验发现,训练数据集的质量对模型结果有重要影响。本次实验从谷歌地图服务中爬取了三个城市的地图数据集用于对比实验,分别是纽约市数据集(通用数据集)、洛杉矶城区数据集、多伦多城区以及周边的遥感影像和相应的

网络地图数据集。

1) 纽约数据集。纽约数据集的层级(缩放级别)为 16 级,其中包含 2 194 对遥感影像及其对应的地图。所有样本都覆盖了纽约市及其周边地区,图像尺寸为 256 × 256。

2) 洛杉矶数据集。同样从谷歌地图服务中收集了 4 791 对美国洛杉矶市区的遥感影像和网络地图,缩放级别为 17 级,空间分辨率为 2.15 m,每幅尺寸为 256 像素 × 256 像素。

3) 多伦多数据集。多伦多城区以及周边地图数据集,缩放级别为 17 级,空间分辨率为 2.15 m。数据集包含 4 828 幅配对遥感影像和网络地图,大小同样为 256 像素 × 256 像素。表 1 列出了以上三个数据集在实验中的使用细节。

##### 4.1.2 评估标准

为量化证明本文提出的半监督遥感影像转换网络地图模型的有效性,在实验中,采用均方根误差、结构相似性、像素准确度三种图像质量评价指标对生成的网络地图进行评估。

表 1 实验数据集统计

Tab. 1 Statistics of experimental dataset

城市	数据集总量	训练样本	配对样本	非配对样本	测试样本	层级	分辨率/m
纽约	2 194	1 560	312	1 248	634	16	4
洛杉矶	4 791	3 300	660	2 640	1491	17	2.15
多伦多	4 828	3 400	680	2 720	1428	17	2.15

1) 均方根误差 (root mean square error, RMSE)。均方误差 (mean square error, MSE) 是指

参数估计值与参数真值之差平方的期望值。MSE 可以评价数据的变化程度, MSE 的值越小, 说明

预测模型描述实验数据具有更好的精确度。RMSE 是均方误差的算术平方根。设预测图像集合为  $\hat{Y} = \{\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_n\}$ , 观察图像集合为  $Y = \{Y_1, Y_2, \dots, Y_n\}$ , RMSE 计算公式为:

$$d_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2} \quad (10)$$

2) 结构相似性 (structural similarity, SSIM)。结构相似性描述了预测图像与观察图像的相似度。分别从亮度、对比度和结构三方面度量图像相似性,取值范围  $[0, 1]$ , SSIM 值越大,表示图像失真越小,计算公式为:

$$d_{\text{SSIM}}(\hat{Y}, Y) = \frac{(2\mu_{\hat{Y}}\mu_Y + c_1)(2\sigma_{\hat{Y}Y} + c_2)}{(\mu_{\hat{Y}}^2 + \mu_Y^2 + c_1)(\sigma_{\hat{Y}}^2 + \sigma_Y^2 + c_2)} \quad (11)$$

其中:  $\mu_{\hat{Y}}$  是  $\hat{Y}$  的均值;  $\mu_Y$  是  $Y$  的均值;  $\sigma_{\hat{Y}Y}$  为  $\hat{Y}$  的方差;  $\sigma_Y$  为  $Y$  的方差;  $c_1$  和  $c_2$  为常数,避免分母接近于 0 时造成的不稳定性。

3) 像素准确度 (pixel accuracy, ACC)。像素准确度计算预测图像的像素精度,给定一张观察图像,其第  $i$  个 RGB 像素值表示为  $(r_i, g_i, b_i)$ , 预测图像对应的第  $i$  个 RGB 像素值表示为  $(r'_i, g'_i, b'_i)$ , 如果  $\max(|r_i - r'_i|, |g_i - g'_i|, |b_i - b'_i|) < \sigma$  (本实验中设置  $\sigma = 5$ ), 则认为预测图像的像素值是准确的<sup>[46]</sup>。

### 4.2 实验设置

本文提出的半监督遥感影像生成网络地图模型基于 Pytorch 深度学习框架,配置了 Intel 酷睿 i7-10700 CPU (2.9 GHz, 8 核, 16 GB RAM) 和单个 NVIDIA GTX 3090 GPU, 24 GB RAM。采用 AdamW 优化器,学习率设为  $10^{-3}$ , 权值衰减为  $10^{-2}$ 。对于这三个数据集的配对数据,批处理大小设置为 4。超参数  $\lambda_{L1}$  和  $\lambda_{\text{TCR}}$  设为 10,  $\lambda_{\text{dice}}$  设为 2。

对于 3.3 节中的几何变换函数  $T_m$ , 本次实验采用顺时针旋转  $90^\circ$ , 对于图像增强设置,采用图 4 中的数据增强方法进行随机组合。

### 4.3 实验结果及评估

为了更全面地验证本文提出的半监督遥感影像生成网络地图模型的有效性,本节以 pix2pix<sup>[10]</sup>、CycleGAN<sup>[17]</sup>、MapGen-GAN<sup>[40]</sup>、SMAPGAN<sup>[42]</sup> 为基线对比模型开展实验分析。其中, pix2pix 和 CycleGAN 为通用的图像域映射模型, MapGen-GAN、SMAPGAN 为专门的遥感影像转换网络地图模型。pix2pix 为全监督模型,必须使用成对的遥感影像和网络地图样本进行训练。

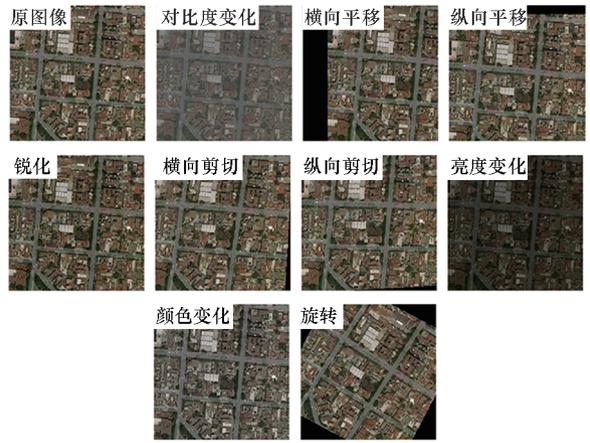


图 4 图像增强方法视觉展示图  
Fig. 4 Visual display diagram of image enhancement methods

CycleGAN 和 MapGen-GAN 为无监督模型, SMAPGAN 为半监督网络地图生成模型。

如表 2 所示,以纽约数据集为例,通过比较生成相同面积的地图的时间,本文模型在快速性上能够满足实时生成的需要,并且和现有经典模型相比,模型仍然能够保持较好的表现。

表 2 不同方法生成 600 km<sup>2</sup> 纽约地图的时间比较  
Tab. 2 Time comparison between different methods for generating a 600 km<sup>2</sup> map based on New York

方法	时间/s
pix2pix	63
CycleGAN	39
本文模型	50

在对比实验中,使用了三种图像质量评价指标来比较本文提出的半监督遥感影像生成网络地图模型与其他基线模型的性能。表 3、表 4、表 5 分别列出了不同模型在纽约数据集、洛杉矶数据集和多伦多数据集上训练的定量结果。pix2pix 作为全监督方法,在表中的得分用斜体表示,性能最佳的用黑体表示。总体上,根据定量实验结果大致可以分析出,本文提出的半监督遥感影像生成网络地图方法在某些指标上略差于 pix2pix。由于 pix2pix 是一种有监督的图像到图像的转换方法,在比较实验中使用了所有配对样本,本文的模型只使用了 20% 的配对数据,因此,在某些评估指标上的得分略低于 pix2pix 是在可接受范围内的。此外,本文模型虽然在纽约数据集和多伦多数据集上有一项指标略低于 SMAPGAN,但总体上是优于其他对比模型的。

表3 不同模型在纽约数据集上的定量对比结果

Tab.3 Quantitative comparison results of different models on New York datasets

方法	RMSE ↓	SSIM ↑	ACC/% ↑
CycleGAN	32.673 1	0.545 3	30.538 5
MapGenGAN	32.054 7	0.564 8	38.644 7
pix2pix	29.018 3	0.594 1	41.566 2
SMAPGAN	30.456 3	<b>0.582 6</b>	39.279 1
本文提出的模型	<b>30.261 7</b>	0.579 4	<b>39.436 8</b>

表4 不同模型在洛杉矶数据集上的定量对比结果

Tab.4 Quantitative comparison results of different models on Los Angeles datasets

方法	RMSE ↓	SSIM ↑	ACC/% ↑
CycleGAN	19.531 7	0.662 4	47.257 3
MapGenGAN	17.742 9	0.702 6	51.791 7
pix2pix	15.618 5	0.738 1	53.842 5
SMAPGAN	16.962 5	0.736 4	55.392 6
本文提出的模型	<b>16.472 8</b>	<b>0.737 9</b>	<b>55.610 4</b>

表5 不同模型在多伦多数据集上的定量对比结果

Tab.5 Quantitative comparison results of different models on Toronto datasets

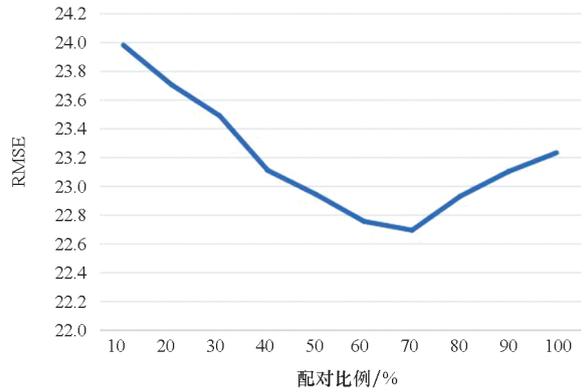
方法	RMSE ↓	SSIM ↑	ACC/% ↑
CycleGAN	28.141 7	0.612 5	44.225 7
MapGenGAN	25.981 3	0.668 3	46.697 2
pix2pix	24.534 2	0.693 5	49.172 8
SMAPGAN	23.821 4	0.672 4	<b>48.954 1</b>
本文提出的模型	<b>23.705 6</b>	<b>0.678 1</b>	48.526 3

具体地,对于纽约数据集,除 pix2pix 外,本文方法在 RMSE 和 ACC 指标上均表现出了最好的性能;对于洛杉矶数据集,在无监督和半监督方法中,本文方法在 3 个指标表现最优,像素准确率达到 55%,甚至超过了 pix2pix;对于多伦多数据集,相比其他无监督/半监督模型,本文方法同样也是在两个指标上得分最高。

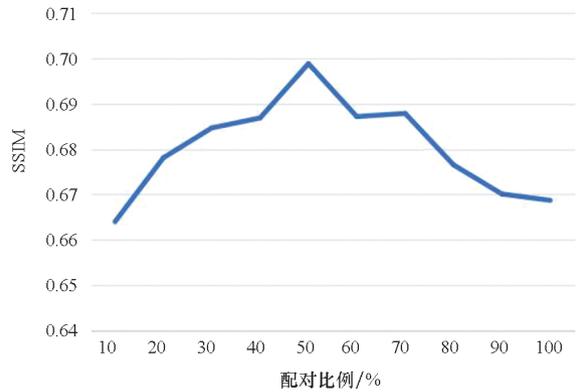
以上结果分析证明了本文提出的半监督遥感影像生成网络地图模型的适用性和有效性,基于变换一致性正则化和样本增强一致性的半监督策略可以有效帮助模型提高生成网络地图的准确度。

实验同时探究了预训练阶段和网络精调阶段对于生成结果的影响。图 5 展示了不同配对比例下,三个定量指标的浮动变化。在精确配对比例较小时,模型性能明显差于全监督模型;但随着比例高于 70%,无监督学习的非配对样本减少,这

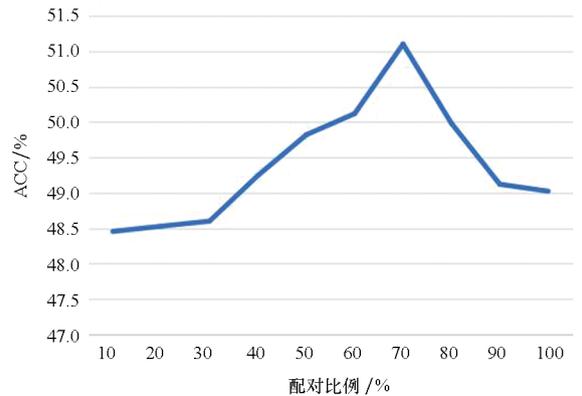
导致数据多样性降低,因此模型的性能反而下降。实验验证了无监督的网络精调阶段能够有效提升模型性能。



(a) 不同配对比例下 RMSE 指标得分  
(a) RMSE scores in the varying paired ratio



(b) 不同配对比例下 SSIM 指标得分  
(b) SSIM scores in the varying paired ratio



(c) 不同配对比例下 ACC 指标得分  
(c) ACC indicator scores in the varying paired ratio

图5 不同配对比例下定量指标在多伦多数据集上的变化  
Fig.5 Fluctuations of the quantitative metrics on the Toronto datasets with varying paired ratio

图 6 展示了不同方法在三种数据集上的定性对比结果。总体上可以看出,本文提出的模型在识别绿地以及生成的道路和建筑物的规整性方面与 pix2pix 不相上下。与另外三种无监督/半监督

基线方法相比,本文提出的模型在区分不同地理特征和填充颜色方面均表现优异。

图 6(a)中的前三幅影像是来自纽约数据集的生成结果,可以明显地看出,本文提出的模型能够较为准确地识别出植被和水体覆盖的区域。图 6(a)中间三幅影像是抽取洛杉矶数据集中的影像,从对应的生成结果可见,本文提出的

模型可以精确地生成建筑物的轮廓,城市地理布局更加规整。图 6(a)中的后三幅影像为多伦多数据集影像,本文模型可以生成准确度和视觉质量较高的网络地图。此外,对于训练样本中相同地理元素的颜色值略有不同的问题,本文提出的模型能够学习接近真实网络地图中的颜色表示。



(a) 遥感影像

(a) Remote sensing imagery



(b) 真值图像

(b) Ground truth



(c) CycleGAN 结果

(c) Results of CycleGAN



(d) MapGen-GAN 结果

(d) Results of MapGen-GAN



(e) SMAPGAN 结果

(e) Results of SMAPGAN



(f) pix2pix 结果

(f) Results of pix2pix



(g) MapGen-GAN 结果  
(g) Results of MapGen-GAN



(h) 本文模型结果  
(h) Results of the proposed model

图 6 不同方法在三个数据集上的定性比较结果

Fig. 6 Qualitative comparison results of different methods on three datasets

### 4.4 消融实验

为了验证模型不同组件的功能,本文开展了消融实验对模型进一步探讨。

本文的半监督模型主要包含三个损失函数:变化一致性正则化( $L_{TCR}$ )、强增强一致性( $L_{dice}$ )和对抗损失函数( $L_{TGAN}$ )。为了讨论不同损失函数的影响,本文首先去除 TCR 损失函数(no-TCR),以此验证变化一致性正则化能否提高几何精度;然后通过去除强增强损失函数来验证数据多样性的增加能够提升模型性能。

表 6~8 展示了消融实验的定量结果。在相同的参数条件下,原始模型的表现优于去除任一损失函数后的模型。通过指标结果,去除强增强一致性损失函数后,模型的性能下降较大,这证明了强增强变化方法,通过提高数据多样性,能够较好地提高模型性能。

表 6 纽约数据集上消融实验的定量对比结果

Tab. 6 Quantitative comparison results of ablation experiments on New York datasets

方法	RMSE ↓	SSIM ↑	ACC/% ↑
原始模型	<b>30.261 7</b>	<b>0.579 4</b>	<b>39.436 8</b>
模型-no- $L_{TCR}$	30.942 6	0.577 4	38.064 8
模型-no- $L_{dice}$	32.475 2	0.568 4	36.368 5

表 7 洛杉矶数据集上消融实验的定量对比结果

Tab. 7 Quantitative comparison results of ablation experiments on Los Angeles datasets

方法	RMSE ↓	SSIM ↑	ACC/% ↑
原始模型	<b>16.472 8</b>	<b>0.737 9</b>	<b>55.610 4</b>
模型-no- $L_{TCR}$	17.984 1	0.702 6	51.791 7
模型-no- $L_{dice}$	19.712 7	0.694 5	50.243 8

表 8 多伦多数据集上消融实验的定量对比结果

Tab. 8 Quantitative comparison results of ablation experiments on Toronto datasets

方法	RMSE ↓	SSIM ↑	ACC/% ↑
原始模型	<b>23.705 6</b>	<b>0.678 1</b>	<b>48.526 3</b>
模型-no- $L_{TCR}$	24.357 5	0.652 3	47.026 8
模型-no- $L_{dice}$	27.854 9	0.620 9	42.915 7

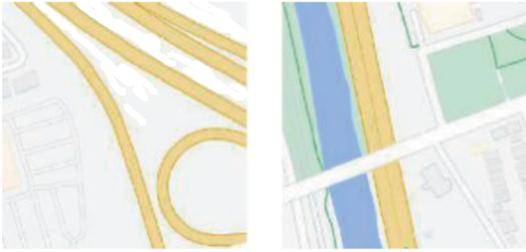
### 4.5 局限性

本文提出的半监督网络地图生成模型旨在利用少量配对样本数据实现高质量遥感影像到网络地图的转换。但是由于数据种类的局限性,以及网络结构固有的能力限制,经过丰富实验验证后发现,在不同地域模型迁移能力和不同地物要素准确识别等方面,模型表现出了一定的局限性。如图 7 所示,在洛杉矶数据集上训练的模型迁移到多伦多数据中,发现实验结果出现了严重的失真。

同时在经过大量实验后,观测到部分数据的参考地图仍然存在要素缺失的问题,经过模型的转换后,这种问题仍然未能较好地解决。经过分析,在模型中施加了变化一致性正则化和强增强



(a) 输入遥感影像  
(a) Remote sensing imagery



(b) 真值图像  
(b) Ground truth



(c) 本文模型  
(c) The proposed model

图 7 由洛杉矶数据集训练模型生成多伦多地图定性结果

Fig. 7 Qualitative results of generating Toronto maps based on the model trained by the Los angeles datasets

一致性损失函数,确保了生成结果的几何准确性,但是生成模型的突出问题就是具有较强的多样性,而这种近似“主观”的特性会造成在地图生成中带来无法控制的结果。如图 8 所示,对于植被区域的绘图就时常被忽略。



(a) 输入遥感影像  
(a) Remote sensing imagery



(b) 真值图像  
(b) Ground truth



(c) 本文模型  
(c) The proposed model

图 8 部分要素缺失的生成地图

Fig. 8 Generate map with missing elements

### 5 结论

综上所述,本文拟面向同一地域不精确匹配的遥感影像与网络地图数据,提出一种基于生成对抗网络的半监督网络地图鲁棒域映射模型,克服非精确匹配数据带来的不一致性问题,设计变换一致性正则化和样本增强一致性学习策略,较全监督/无监督学习方法能表现出更好的模型泛化性能。本文的主要贡献包括三个方面:

1) 提出了一种新颖的基于生成对抗网络 GAN 的半监督遥感影像生成网络地图模型,该模型解决了现有的深度学习网络地图域映射方法所面临的训练数据不精确匹配导致模型泛化性能低的问题。

2) 设计了一致性正则化的半监督训练策略。首先使用小部分精确匹配的数据预训练网络,然后引入多种影像增强方法使其产生几何变化,进而先验知识的基础上使用大量的非配对数据以无监督方式训练网络。最后采用变换一致性损失函数和样本增强一致性损失函数保证生成网络地图的几何精度。

3) 基于三个不同地域的数据集开展了充分的实验验证,定量和定性的实验结果表明本文提出的模型明显优于其他半监督或无监督模型,甚至性能接近于全监督模型。

未来的工作拟针对海量易获取的非精确匹配的遥感影像和网络地图样本,设计专门的自动化数据清洗方法,解决人工筛选成本高的问题。此外,还将研究在不成完整金字塔的情况下,保证不同层级网络地图中地物表述的一致性问题。

### 参考文献 (References)

[1] 张兵. 遥感大数据时代与智能信息提取[J]. 武汉大学学报(信息科学版), 2018, 43(12): 1861-1871.

- ZHANG B. Remotely sensed big data era and intelligent information extraction[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1861 - 1871. (in Chinese)
- [2] 艾廷华. 深度学习赋能地图制图的若干思考[J]. *测绘学报*, 2021, 50(9): 1170 - 1182.
- AI T H. Some thoughts on deep learning enabling cartography[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(9): 1170 - 1182. (in Chinese)
- [3] 付莹, 梁帅哲. 一种基于遥感影像的网络地图智能生成方法及系统:CN202110269420. X[P]. 2022-09-06.
- FU Y, LIANG S Z. A method and system for intelligent generation of network maps based on remote sensing images: CN202110269420. X[P]. 2022-09-06. (in Chinese)
- [4] LI Z L, HUANG P Z. Quantitative measures for spatial information of maps[J]. *International Journal of Geographical Information Science*, 2002, 16(7): 699 - 709.
- [5] LIANG J Y, LUGMAYR A, ZHANG K, et al. Hierarchical conditional flow: a unified framework for image super-resolution and image rescaling [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 4056 - 4065.
- [6] ZHANG Y, LI D S, LAW K L, et al. IDR: self-supervised image denoising via iterative data refinement [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 2088 - 2097.
- [7] ZHENG A L, ZHANG Y A, ZHANG X Y, et al. Progressive end-to-end object detection in crowded scenes [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 847 - 856.
- [8] TEE W Z, DAVE R, SELIYA J, et al. A close look into human activity recognition models using deep learning[C]//Proceedings of the 3rd International Conference on Computing, Networks and Internet of Things (CNIOT), 2022: 201 - 206.
- [9] HAN L G, MIN M R, STATHOPOULOS A, et al. Dual projection generative adversarial networks for conditional image generation[EB/OL]. (2021 - 11 - 29) [2024 - 01 - 05]. <https://arxiv.org/abs/2108.09016v2>.
- [10] WU C H, DE LA TORRE F. A latent space of stochastic diffusion models for zero-shot image editing and guidance[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 7344 - 7353.
- [11] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5967 - 5976.
- [12] HOSSEINPOUR H, SAMADZADEGAN F, JAVAN F D. CMGFNet: a deep cross-modal gated fusion network for building extraction from very high-resolution remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 184: 96 - 115.
- [13] ZHANG C, JIANG W S, ZHANG Y, et al. Transformer and CNN hybrid deep neural network for semantic segmentation of very-high-resolution remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 4408820.
- [14] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139 - 144.
- [15] KINGMA D P, WELLING M. Stochastic gradient VB and the variational auto-encoder[EB/OL]. (2022 - 12 - 10) [2024 - 01 - 07]. <https://arxiv.org/abs/1312.6114v7>.
- [16] HO J, JAIN A N, ABBEEL P. Denoising diffusion probabilistic models[C]//Proceedings of Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 2020.
- [17] MIRZA M, OSINDERO S. Conditional generative adversarial nets [EB/OL]. (2014 - 11 - 06) [2024 - 01 - 12]. <https://arxiv.org/abs/1411.1784>.
- [18] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: 2242 - 2251.
- [19] WU H K, ZHENG S, ZHANG J G, et al. GP-GAN: towards realistic high-resolution image blending[C]//Proceedings of the 27th ACM International Conference on Multimedia, 2019: 2487 - 2495.
- [20] DEMIR U, UNAL G. Patch-based image inpainting with generative adversarial networks[EB/OL]. (2018 - 03 - 20) [2024 - 02 - 07]. <https://arxiv.org/abs/1803.07422>.
- [21] KIM H, CHOI Y, KIM J, et al. Exploiting spatial dimensions of latent in GAN for real-time image editing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 852 - 861.
- [22] MENG C L, HE Y T, SONG Y, et al. SDEdit: guided image synthesis and editing with stochastic differential equations[EB/OL]. (2022 - 01 - 05) [2024 - 02 - 07]. <https://arxiv.org/abs/2108.01073v2>.
- [23] NGUYEN T, OJHA U, LI Y H, et al. Edit one for all: interactive batch image editing[EB/OL]. (2024 - 01 - 18) [2024 - 03 - 18]. <https://arxiv.org/abs/2401.10219v1>.
- [24] LI B C, LI X, ZHU H X, et al. SeD: semantic-aware discriminator for image super-resolution[EB/OL]. (2024 - 02 - 29) [2024 - 03 - 20]. <https://arxiv.org/abs/2402.19387v1>.
- [25] SOLANO-CARRILLO E, RODRIGUEZ A B, CARRILLO-PEREZ B, et al. Look ATME: the discriminator mean entropy needs attention[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2023: 787 - 796.
- [26] XU Y R, PARK T, ZHANG R, et al. VideoGigaGAN: towards detail-rich video super-resolution[EB/OL]. (2024 - 03 - 01) [2024 - 03 - 20]. <https://arxiv.org/abs/2404.12388v2>.
- [27] ZHANG K H, REN W Q, LUO W H, et al. Deep image deblurring: a survey[EB/OL]. (2022 - 05 - 28) [2024 - 03 - 20]. <https://arxiv.org/abs/2201.10700>.
- [28] ZHONG Y, LIU L Z, ZHAO D, et al. A generative adversarial network for image denoising[J]. *Multimedia Tools and Applications*, 2020, 79(23): 16517 - 16529.
- [29] VAN DEN OORD A, KALCHBRENNER N, KAVUKCUOGLU K. Pixel recurrent neural networks[C]//Proceedings of the 33rd International Conference on Machine Learning, 2016: 1747 - 1756.
- [30] DINH L, SOHL-DICKSTEIN J, BENGIO S. Density estimation using real NVP[EB/OL]. (2017 - 02 - 27) [2024 - 03 - 20]. <https://arxiv.org/abs/1605.08803>.
- [31] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes

- dataset for semantic urban scene understanding [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 3213 – 3223.
- [32] RICHARDSON E, ALALUF Y, PATASHNIK O, et al. Encoding in style: a StyleGAN encoder for image-to-image translation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 2287 – 2296.
- [33] GANGUL S, GARZON P, GLASER N. GeoGAN: a conditional GAN with reconstruction and style loss to generate standard layer of maps from satellite images[EB/OL]. (2019 – 04 – 25) [2024 – 03 – 21]. <https://arxiv.org/abs/1902.05611>.
- [34] WANG T C, LIU M Y, ZHU J Y, et al. High-resolution image synthesis and semantic manipulation with conditional GANs[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8798 – 8807.
- [35] YOO D, KIM N, PARK S, et al. Pixel-level domain transfer[C]//Proceedings of the European Conference on Computer Vision, 2016.
- [36] KIM T, CHA M, KIM H, et al. Learning to discover cross-domain relations with generative adversarial networks[C]// Proceedings of the 34th International Conference on Machine Learning, 2017: 1857 – 1865.
- [37] LIU M Y, BREUEL T, KAUTZ J. Unsupervised image-to-image translation networks [EB/OL]. (2018 – 07 – 23) [2024 – 03 – 22]. <https://arxiv.org/abs/1703.00848>.
- [38] 张刚. 基于深度学习的遥感图像语义分割关键技术研究[D]. 成都: 中国科学院大学(中国科学院光电技术研究所), 2020.
- ZHANG G. Research on key technologies of remote sensing image semantic segmentation based on deep learning [D]. Chengdu: Institute of Optics and Electronics, Chinese Academy of Sciences, 2020. (in Chinese)
- [39] KANG Y H, GAO S, ROTH R E. Transferring multiscale map styles using generative adversarial networks [J]. International Journal of Cartography, 2019, 5(2/3): 115 – 141.
- [40] SONG J Q, LI J, CHEN H, et al. MapGen-GAN: a fast translator for remote sensing image to map via unsupervised adversarial learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 2341 – 2357.
- [41] SONG J Q, LI J, CHEN H, et al. RSMT: a remote sensing image-to-map translation model via adversarial deep transfer learning[J]. Remote Sensing, 2022, 14(4): 919.
- [42] CHEN X, CHEN S Q, XU T, et al. SMAPGAN: generative adversarial network-based semisupervised styled map tile generation method[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(5): 4388 – 4406.
- [43] DING Z P, WANG B, SUN S F, et al. Improved SmapGAN remote sensing image map generation based on multi-head self-attention and carafe [J]. Journal of Applied Remote Sensing, 2024, 18(1): 014526.
- [44] FU Y, LIANG S Z, CHEN D D, et al. Translation of aerial image into digital map via discriminative segmentation and creative generation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1 – 15.
- [45] CHEN L W, FANG Z, FU Y. Consistency-aware map generation at multiple zoom levels using aerial image [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 5953 – 5966.
- [46] VAN ENGELEN J E, HOOS H H. A survey on semi-supervised learning[J]. Machine Learning, 2020, 109(2): 373 – 440.