

基于空间分离表征 GAN 的遥感图像模式互转

韩子硕, 王春平*, 付强, 赵斌

陆军工程大学石家庄校区电子与光学工程系, 河北 石家庄 050003

摘要 针对合成孔径雷达图像与光学遥感图像模式差异大、相互转换困难的问题, 基于现有空间分离图像转换框架, 提出了一种基于空间分离表征的循环一致性生成对抗网络(GAN)。以更深的网络层和跳跃连接完成图像风格和内容分离, 通过学习内容映射关系, 完成内容特征转换, 而后组合目标风格特性实现图像转换。利用 PatchGAN 判别器, 强化模型的图像细节信息生成能力, 并新增目标误差损失和生成重建损失将转换任务限制为一对一映射, 减少信息添加, 约束生成网络。在 SEN1-2、SARptical、WHU-SEN-City 数据集上进行实验验证, 相较于其他图像转换算法, 所提方法能够有效实现两类遥感图像互转, 生成图像清晰度高、细节特征完整、真实感强。

关键词 遥感; 图像转换; 合成孔径雷达; 光学遥感图像; 循环一致性生成对抗网络

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202141.0728003

Remote Sensing Image Mode Translation by
Spatial Disentangled Representation Based GAN

Han Zishuo, Wang Chunping*, Fu Qiang, Zhao Bin

Department of Electronic and Optical Engineering, Shijiazhuang Campus, Army Engineering University,
Shijiazhuang, Hebei 050003, China

Abstract Resting on the translation framework of spatially separated images, we proposed a cycle-consistent generative adversarial network (GAN) based on spatial disentangled representation to address the large mode difference and difficult translation between synthetic aperture radar images and optical remote sensing images. The proposed model separates images into style and content features by a deeper network layer and jump connection. Furthermore, the content features are translated by content mapping learning and combined with target style features for image translation. In addition, PatchGAN, as the discriminator, enhances the image detail generation, and target error loss and generation & reconstruction loss are introduced to limit the translation task to one-to-one mapping, thus reducing the information added and constraining the GAN. The experimental results in SEN1-2, SARptical, and WHU-SEN-City datasets show that compared with other image translation algorithms, the proposed method can translate two types of remote sensing images and generate images of high resolution, complete detail features, and strong authenticity.

Key words remote sensing; image translation; synthetic aperture radar; optical remote sensing image; cycle-consistent adversarial networks

OCIS codes 280.4750; 100.4994; 100.4996; 150.1135

1 引言

合成孔径雷达(SAR)和光学遥感(ORS)探测

器作为两项尖端的对地探测遥感技术, 已被广泛应用于地图绘制、灾难预测、环境监控、海洋资源利用、目标检测等多个方面^[1-3]。SAR 可以全天时、全天

收稿日期: 2020-09-27 修回日期: 2020-11-13; 录用日期: 2020-11-24

基金项目: 军内科研项目(LJ20191A040155)

* E-mail: wang_c_p@163.com

候地实施对地探测任务,能够实现高分辨率成像^[4]。但由于其独特的成像机制和复杂的电磁波散射过程,导致所采集图像含有大量相干斑噪声,且地物信息抽象、可读性差,非专业人员根本无法进行解译。而 ORS 图像正好与之相反,其色彩鲜明的地物呈现和高清晰的图像质量,非常符合人类视觉体系读取,但恶劣的天气和云层会造成 ORS 图像部分缺失。因此实现两者之间的互转,可弥补各自不足,对地表探测和研究具有重大意义。

近几年,随着深度学习技术的迅速崛起,越来越多的网络模型被应用于图像处理领域^[5-8]。生成对抗网络(GAN)^[9]作为一种新的生成模型,在图像生成和转换、信息安全、视觉计算等领域展现出巨大的应用价值。许多基于 GAN 的图像转换模型取得了巨大成功。Pix2pix^[10]可实现语义图到真实场景图、卫星图到地图、日景图到夜景图的转换。循环一致性生成对抗网络(Cycle-GAN)^[11-12]可通过训练非成对图像,实现场景图四季转换、照片与油画互转、红外到可见光人脸图像转换、图像中动物种类的转换。对偶生成对抗网络(Dual-GAN)^[13]可实现素描到照片、中国画到油画、图像中物体材质之间的转换。以上 GAN 模型虽可以完成一些特定场景转换,但难以胜任模式相差较大的转换任务。基于此,空间分离表征方法被引至 GAN 中。DRIT^[14]将图像分解至内容空间和特性空间,利用所提取的内容特征和属性向量完成转换任务。MUNIT^[15]将图像分离成内容特征和风格特征,通过为内容特征赋予不同的风格特性实现多模态图像转换。在此基础上,DSMAP^[16]共享潜在的内容空间,并建立内容特征映射关系,使转换后的内容特征能够与目标域特征保持高度匹配。但以上方法均面向一对多的图像转换任务,并不适用于 SAR 图像与 ORS 图像互转中所要求的严格对应关系。

基于以上分析,针对 SAR 图像与 ORS 图像互转问题,本文提出了基于空间分离表征的循环一致性生成对抗网络(SDR-CycleGAN)。该方法在 DSMAP 基础之上,利用更深的网络层和跳跃连接,充分挖掘分离空间中的内容特征信息,加强浅层与深层之间的信息传递,建立高匹配度的内容特征映射,利用 PatchGAN^[10]构建判别器,进一步优化生成图像细节信息,并引入目标误差损失和生成重构损失,将循环网络限制为一对一映射,减少无用计算

和错误信息添加的同时,有效避免了梯度消失和模式坍塌,稳定了网络训练过程。基于实测数据的实验结果证明,所提网络模型可有效提升 SAR 图像与 ORS 图像之间互转的准确性和鲁棒性。

2 基于空间分离表征的循环一致性生成对抗网络

2.1 网络结构

SDR-CycleGAN 的基本框架如图 1 所示,其包括两个风格编码器 E_A^s 和 E_B^s ,两个内容编码器 E_A^c 和 E_B^c ,两个解码器 DE_A 和 DE_B ,两个判别器 D_A 和 D_B ,两个多层感知器(MLP),以及由共享残差块和两个内容空间映射组成的内容特征转换模块。两个 E^s 模型相同,其他成对模块同理。对于 SAR 图像至 ORS 图像转换,即 A 域至 B 域的转换(A→B),其过程描述如下。

目标 $x_A \rightarrow x_{A \rightarrow B}$;

输入 SAR 图像 x_A 和 ORS 图像 x_B ;

步骤 1 提取 x_A 内容特征 $c_A = E_A^c(x_A)$,提取 x_B 风格特征 $s_B = E_B^s(x_B)$;

步骤 2 c_A 经共享残差块 E_{res}^c 和 A2B 内容映射 $\Phi_{c \rightarrow c_B}$,完成 A→B 内容转换: $c_A \rightarrow c'_A = E_{res}^c(c_A) \rightarrow c'_{A \rightarrow B} = \Phi_{c \rightarrow c_B}(c'_A)$;

步骤 3 $c'_{A \rightarrow B}$ 经解码器 DE_B 完成最终转换: $c'_{A \rightarrow B} \rightarrow x_{A \rightarrow B} = DE_B(c'_{A \rightarrow B}, MLP(s_B))$;

输出 $x_{A \rightarrow B}$ 。

ORS 图像至 SAR 图像转换(即 B→A)同 A→B 过程类似。将 $x_{A \rightarrow B}$ 作为 ORS 图像输入 B→A 转换网络,可得到 x_A 的重构图像 $x_{A \rightarrow B \rightarrow A}$,完成 $x_A \rightarrow x_{A \rightarrow B} \rightarrow x_{A \rightarrow B \rightarrow A}$ 循环。

判别器 D_A 用以判定生成 SAR 图像 $x_{B \rightarrow A}$ 的真实性,判别器 D_B 用以判定生成 ORS 图像 $x_{A \rightarrow B}$ 的真实性。训练过程中,在 CycleGAN 生成对抗损失和循环一致损失基础之上,附加内容和风格重构损失、图像重构损失、内容空间自映射损失、生成图像与目标图像之间的 L_1 目标误差损失,以及生成重构损失,控制网络走向,稳定训练过程。

2.2 风格编码器和内容编码器

风格编码器用以提取源域和目标域各自的风格特征(s_A/s_B),网络结构如图 2 所示,由六个子风格编码器(S-Encode_*)和一个全局平均池化层组成,其中,C 表示卷积,R 表示修正线性单元(ReLU)。

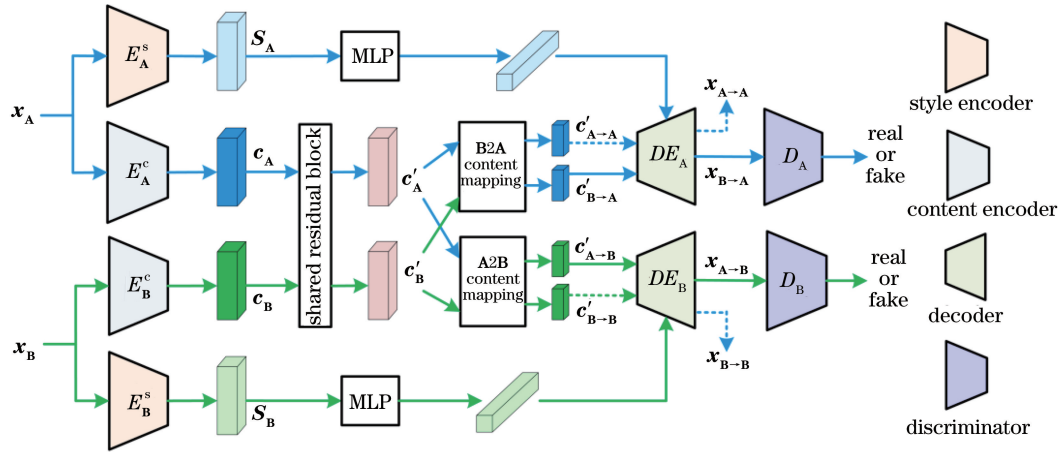


图 1 SDR-CycleGAN 的基本框架

Fig. 1 Basic framework of SDR-CycleGAN

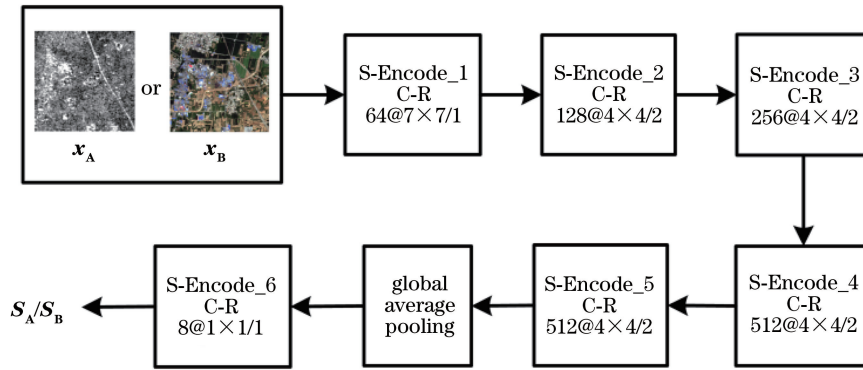


图 2 风格编码器网络结构

Fig. 2 Network structure of style encoder

内容编码器网络结构如图 3 所示,由 13 个子内容编码器(C-Encode_*)组成,其中,IN 表示实例归一化,Add 表示特征相加。内容编码器用以提取源

域和目标域各自的内容特征(c_A/c_B),同时由 C-Encode_3 和 C-Encode_4 输出各域的前期内容特征($pre_c_A^1/pre_c_B^1$ 和 $pre_c_A^2/pre_c_B^2$)。

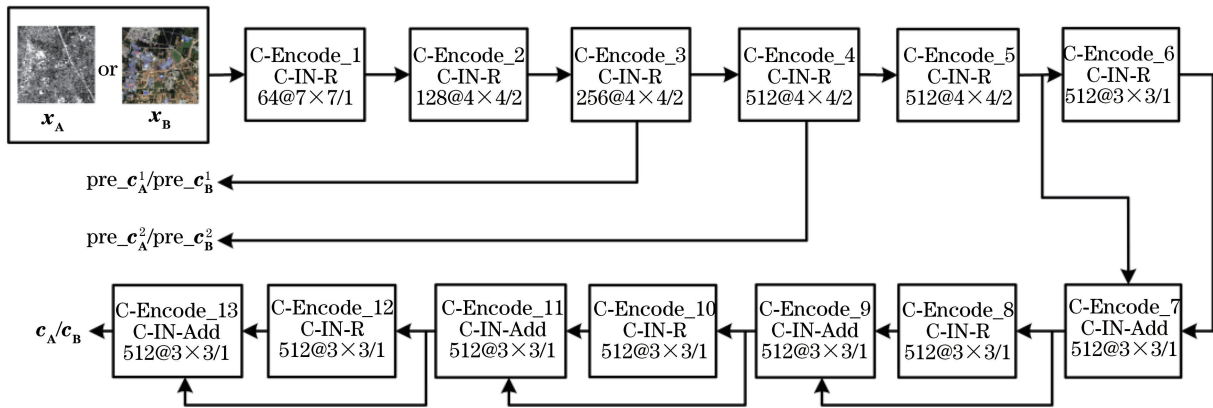


图 3 内容编码器网络结构

Fig. 3 Network structure of content encoder

2.3 内容域转换模块

内容域转换模块用以完成源域与目标域之间的内容特征互转,网络结构如图 4 所示,由一个共享残

差块(Shared-Res_*)和两个内容映射分支(C-Map-A/B_*)组成,其中,LR 表示带泄露修正线性单元(LeakyReLU),DC 表示反卷积,LN 表示层归

一化。两个内容映射分支的第 1、2 层通过跳跃连接融合内容编码器输出的前期特征,以便充分利用卷积层中的内容特征信息。C-Map-A 内容映射分支

输出内容重构特征 $c'_{A \rightarrow A}$ 和内容映射特征 $c'_{B \rightarrow A}$, C-Map-B 内容映射分支输出 $c'_{A \rightarrow B}$ 和 $c'_{B \rightarrow B}$ 。

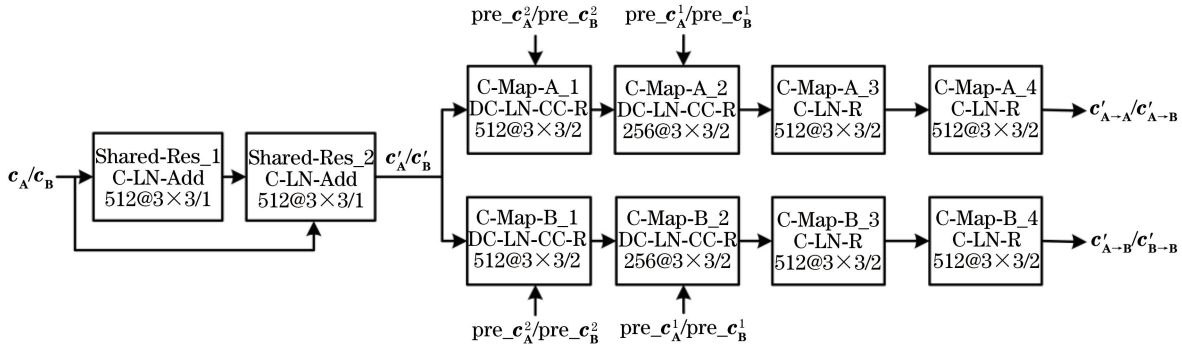


图 4 内容域转换模块网络结构

Fig. 4 Network structure of content domain translation module

2.4 解码器

解码器用以整合内容特征和风格特征,完成最终图像转换,网络结构如图 5 所示,由 MLP 分支与解码器(Decode_*)分支组成,其中,FC 表示全连接层,AdaIN^[17]表示自适应实例归一化。MLP 以

风格特征 s_A 或 s_B 为输入,输出 AdaIN 所需参数。解码器分支最后一层使用双曲正切函数(Tanh)作为激活函数,输出重构图像 $x_{A \rightarrow A}/x_{B \rightarrow B}$ 和转换图像 $x_{A \rightarrow B}/x_{B \rightarrow A}$ 。

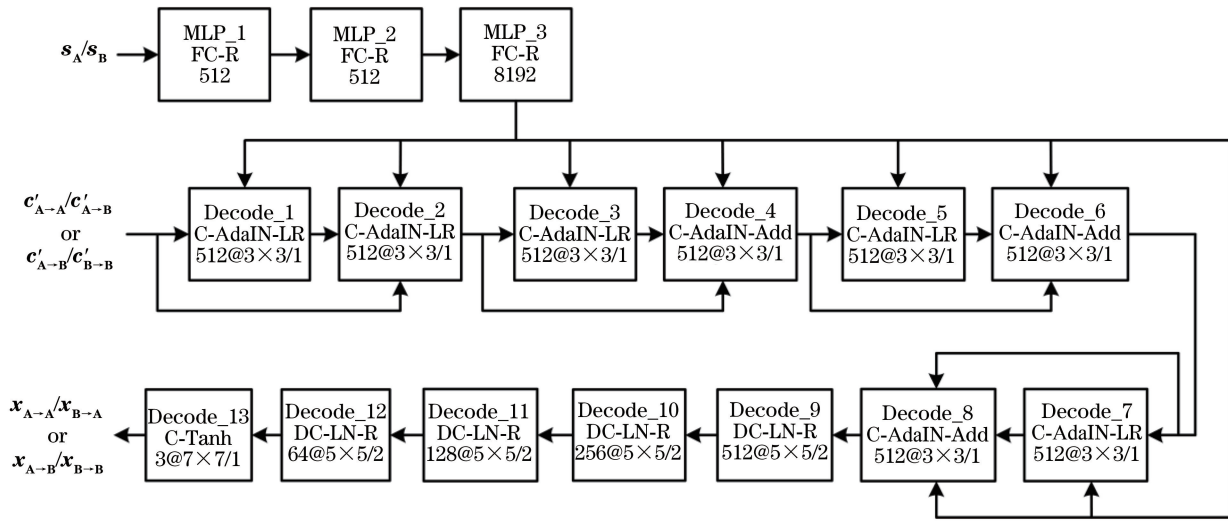


图 5 解码器网络结构

Fig. 5 Network structure of decoder

2.5 判别器

传统 GAN 的判别器由一系列的卷积层组成,输出一个介于 0~1 之间的标量,即输入样本判定为真的概率。其输出结果可以看作是一个整体图像的加权值,无法体现局部图像特征。由于 SAR 图像与 ORS 图像模式差异大,此种结构无法满足高精度要求。因此,本文利用 70×70 的 PatchGAN^[10] 构建判别器,如图 6 所示,其中 B 表示批量归一化,S 表示 Sigmoid 函数。判别器的工作过程如图 7 所示,其输入 $256 \text{ pixel} \times 256 \text{ pixel}$ 大小的图像,输出一个

30×30 的矩阵,每个矩阵元素对应输入图像中一个 70×70 的 Patch,代表此 Patch 为真的概率,最终取所有元素的均值作为整幅生成图像的判定。值得注意的是,PatchGAN 处理的并不是图像的裁剪块,而是整幅图像,判定矩阵中的每个元素实际上代表原图中一个 70×70 的感受野,因此称为 70×70 PatchGAN。这种方法通过对每一个 Patch 进行差别判定,实现了局部图像特征的提取和表征,有利于实现特征差异较大的图像转换。

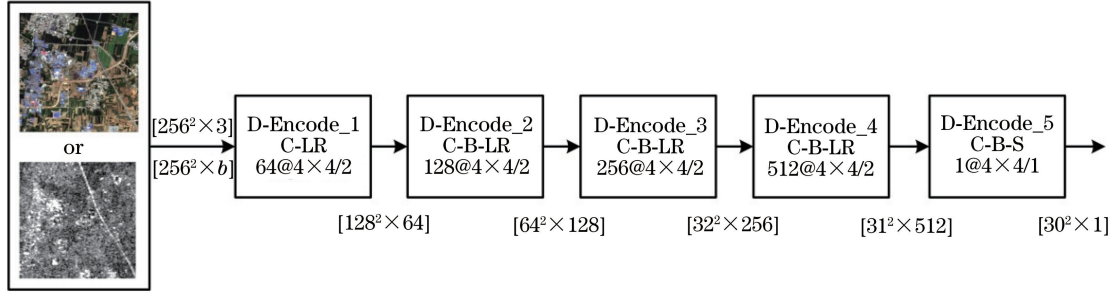


图 6 判别器网络结构

Fig. 6 Network structure of discriminator

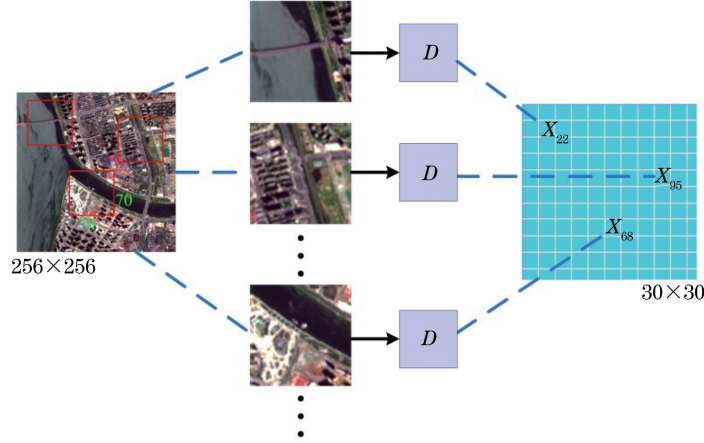


图 7 判别器工作过程示意图

Fig. 7 Working process diagram of discriminator

3 损失函数

本文联合对抗损失、循环一致性损失、内容空间自映射损失、内容重构损失、风格重构损失、图像重构损失、生成重构损失和目标误差损失对网络进行训练。按照循环模型 $A \rightarrow B \rightarrow A$ 和 $B \rightarrow A \rightarrow B$ 两个路径划分,前六项损失属于单路径损失,后两项损失属于跨路径损失。

3.1 对抗损失和循环一致性损失

传统 GAN 中使用的交叉熵损失往往会导致位于决策边界正确一侧但仍远离真实数据的样本梯度迅速消失。最小二乘损失可以对距离决策边界较远的样本(无论其是否被正确分类)进行惩罚,从而缓解梯度消失问题。因此,SDR-CycleGAN 利用最小二乘损失函数构建对抗损失。

对于 B 域至 A 域转换过程和判别器 D_A ,对抗损失表示为

$$L_{D_{adv}}^A = \frac{1}{2} \mathbb{E}_{x_A} \{ [D_A(x_A) - 1]^2 \} + \frac{1}{2} \mathbb{E}_{x_{B \rightarrow A}} \{ [D_A(x_{B \rightarrow A})]^2 \}, \quad (1)$$

$$L_{G_{adv}}^A = \frac{1}{2} \mathbb{E}_{x_{B \rightarrow A}} \{ [D_A(x_{B \rightarrow A}) - 1]^2 \}, \quad (2)$$

式中: $\mathbb{E}(\cdot)$ 表示求取期望; $D(\cdot)$ 表示输入图像为真的概率; x_A 为来自 A 域的图像; $x_{B \rightarrow A}$ 为由 B 域转换至 A 域的图像。

同理,对于 A 域至 B 域转换过程和判别器 D_B ,对抗损失表示为

$$L_{D_{adv}}^B = \frac{1}{2} \mathbb{E}_{x_B} \{ [D_B(x_B) - 1]^2 \} + \frac{1}{2} \mathbb{E}_{x_{A \rightarrow B}} \{ [D_B(x_{A \rightarrow B})]^2 \}, \quad (3)$$

$$L_{G_{adv}}^B = \frac{1}{2} \mathbb{E}_{x_{A \rightarrow B}} \{ [D_B(x_{A \rightarrow B}) - 1]^2 \}, \quad (4)$$

式中: x_B 为来自 B 域的图像; $x_{A \rightarrow B}$ 为由 A 域转换至 B 域的图像。

对于 $A \rightarrow B \rightarrow A$ 和 $B \rightarrow A \rightarrow B$ 两个转换过程,使用 L_1 范数损失构建循环一致性损失,

$$L_{cyc} = \mathbb{E}_{x_A} [\|x_{A \rightarrow B \rightarrow A} - x_A\|_1] + \mathbb{E}_{x_B} [\|x_{B \rightarrow A \rightarrow B} - x_B\|_1], \quad (5)$$

式中: $\|\cdot\|_1$ 为 L_1 范数; $x_{A \rightarrow B \rightarrow A}$ 为 x_A 的重构图像; $x_{B \rightarrow A \rightarrow B}$ 为 x_B 的重构图像。

3.2 内容空间自映射损失

共享残差块 E_{res}^c 和内容空间映射 $\Phi_{c \rightarrow c_A} / \Phi_{c \rightarrow c_B}$ 用以完成 $\mathbf{A} \leftrightarrow \mathbf{B}$ 之间内容特征互转, 同时计算 \mathbf{A} 域与 \mathbf{B} 域内容特征自映射, 即: $c_{A \rightarrow A} / c_{B \rightarrow B}$ 。 $\mathbf{A} \rightarrow \mathbf{A}$ 内容特征自映射过程可表示为 $c_A \rightarrow c'_A = E_{res}^c(c_A) \rightarrow c'_{A \rightarrow A} = \Phi_{c \rightarrow c_A}(c'_A)$ 。高效的内容特征转换要求 $\Phi_{c \rightarrow c_A}$ 的输出具有 \mathbf{A} 域独特的内容特性。因此, 利用 $c'_{A \rightarrow A}$ 与 c_A 之间的 L_1 范数误差损失约束内容特征转换过程。 $\mathbf{B} \rightarrow \mathbf{B}$ 内容特征自映射过程同理。内容空间自映射损失表示为

$$L_{dsc} = \mathbb{E}_{c_A} \{ \| c_A - \Phi_{c \rightarrow c_A} [E_{res}^c(c_A)] \|_1 \} + \mathbb{E}_{c_B} \{ \| c_B - \Phi_{c \rightarrow c_B} [E_{res}^c(c_B)] \|_1 \}, \quad (6)$$

式中: $E_{res}^c(\cdot)$ 表示共享残差函数; $\Phi(\cdot)$ 表示内容特征转换函数; $c_{A/B}$ 表示 \mathbf{A} 域或 \mathbf{B} 域的内容特征。

3.3 内容重构损失、风格重构损失和图像重构损失

内容重构损失、风格重构损失和图像重构损失与循环一致性损失的目的一样, 均用以确保循环网络在单路径上可以高效完成图像转换和重构。

内容重构损失定义为

$$L_c = \mathbb{E}_{x_A, x_{A \rightarrow B}} \{ \| E_B^c [E_{res}^c(x_{A \rightarrow B})] - E_A^c [E_{res}^c(x_A)] \|_1 \} + \mathbb{E}_{x_B, x_{B \rightarrow A}} \{ \| E_A^c [E_{res}^c(x_{B \rightarrow A})] - E_B^c [E_{res}^c(x_B)] \|_1 \}, \quad (7)$$

式中: $E^c(\cdot)$ 表示内容编码函数。

风格重构损失定义为

$$L_s = \mathbb{E}_{c_A, x_{B \rightarrow A}} [\| E_A^s(x_{B \rightarrow A}) - c_A \|_1] + \mathbb{E}_{c_B, x_{A \rightarrow B}} [\| E_B^s(x_{A \rightarrow B}) - c_B \|_1], \quad (8)$$

式中: $E^s(\cdot)$ 表示风格编码函数; $s_{A/B}$ 表示 \mathbf{A} 域或 \mathbf{B} 域的风格特征。

图像重构损失定义为

$$L_x = \mathbb{E}_{x_A, x_{A \rightarrow A}} [\| x_{A \rightarrow A} - x_A \|_1] + \mathbb{E}_{x_B, x_{B \rightarrow B}} [\| x_{B \rightarrow B} - x_B \|_1]。 \quad (9)$$

3.4 生成重构损失和目标误差损失

对于循环网络两个路径, x_A 、 $x_{B \rightarrow A}$ 和 $x_{A \rightarrow B \rightarrow A}$ 同属于 \mathbf{A} 域图像, 理想情况下, 三者应保持一致。 x_B 、

$x_{A \rightarrow B}$ 和 $x_{B \rightarrow A \rightarrow B}$ 同理。鉴于此, 我们在循环网络两个路径之间建立损失函数通道, 增加生成重构损失和目标误差损失, 将生成图像与另一路重建图像和输入图像分别进行比较, 以此来进一步约束整个图像生成过程, 以便利用 GAN 的学习能力减少错误映射和无用信息。

生成重构损失定义为

$$L_{gre} = \mathbb{E}_{x_{A \rightarrow B \rightarrow A}, x_{B \rightarrow A}} [\| x_{A \rightarrow B \rightarrow A} - x_{B \rightarrow A} \|_1] + \mathbb{E}_{x_{B \rightarrow A \rightarrow B}, x_{A \rightarrow B}} [\| x_{B \rightarrow A \rightarrow B} - x_{A \rightarrow B} \|_1]。 \quad (10)$$

目标误差损失定义为

$$L_p = \mathbb{E}_{x_A, x_{B \rightarrow A}} [\| x_{B \rightarrow A} - x_A \|_1] + \mathbb{E}_{x_B, x_{A \rightarrow B}} [\| x_{A \rightarrow B} - x_B \|_1]。 \quad (11)$$

此两项损失要求网络输入与输出相互对应, 因此需要对成对样本对网络进行训练, 而此情况也正适用于 SAR 图像与 ORS 图像之间互转的唯一性。

3.5 总损失函数

将各项损失按照一定比例相加, 构成 SDR-CycleGAN 的整体损失函数, 表示为

$$L_{G_{total}} = \lambda_{adv} (L_{G_{adv}}^A + L_{G_{adv}}^B) + \lambda_{cyc} L_{cyc} + \lambda_{dsc} L_{dsc} + \lambda_c L_c + \lambda_s L_s + \lambda_x L_x + \lambda_p L_p + \lambda_{gre} L_{gre}, \quad (12)$$

$$L_{D_{total}} = \lambda_{adv} (L_{D_{adv}}^A + L_{D_{adv}}^B), \quad (13)$$

式中: λ 是调整网络各项损失占比的参数。

4 实验设置与评估参数

4.1 数据集

为了验证本文所提模型在 SAR 图像与 ORS 图像互转中的有效性, 采用 SEN1-2 数据集^[18]、SARptical 数据集^[19]和 WHU-SEN-City 数据集^[20]作为实验对象。

SEN1-2 数据集包含了从 Sentinel-1 和 Sentinel-2 场景中提取的 282384 对对应的 SAR 和 ORS 图像, 这些图像涵盖全球各地和所有气象季节。数据集中的图像大小为 256 pixel \times 256 pixel, 样例如图 8 所示。

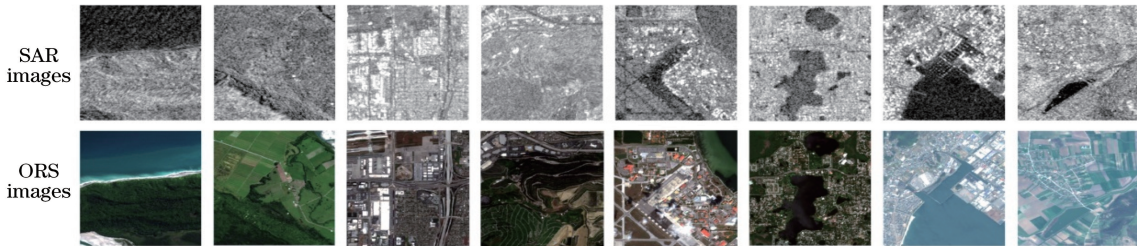


图 8 SEN1-2 数据集样例

Fig. 8 Samples of SEN1-2 dataset

SAR_{ptical} 数据集由 10108 对对应的高分辨率 TerraSAR-X 图像和柏林市中心的航空 UltraCAM 光学图像组成。该数据集中的图像大小为 112 pixel×112 pixel,在输入网络之前,首先将其扩张为 256 pixel×256 pixel 大小。SAR_{ptical} 数据集样例如图 9 所示。

WHU-SEN-City 数据集包含 34 对对应的 SAR 场景图与 ORS 场景图,数据覆盖中国 32 个

城市。SAR 图像采自多波段哨兵-1 号遥感卫星, ORS 图像采自高分辨率多光谱哨兵-2 号遥感卫星。数据集样例如图 10 所示。网络训练之前,将数据集中的场景图以 25% 的覆盖率剪切成 11761 对大小为 256 pixel×256 pixel 的图像块,图像块样例如图 11 所示。场景图剪切过程如图 12 所示,重叠区域宽度由覆盖率乘以图像块宽度确定。

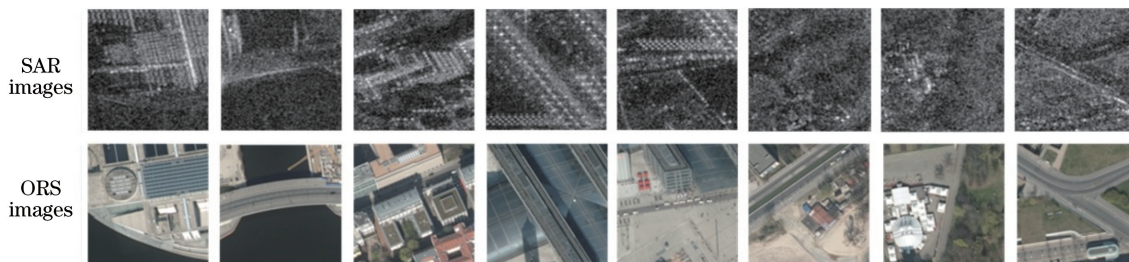


图 9 SAR_{ptical} 数据集样例
Fig. 9 Samples of SAR_{ptical} dataset

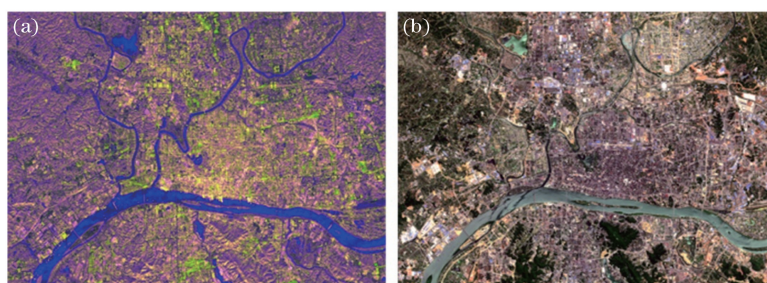


图 10 WHU-SEN-City 数据集样例。(a) SAR 图像;(b) ORS 图像
Fig. 10 Samples of WHU-SEN-City dataset. (a) SAR image; (b) ORS image

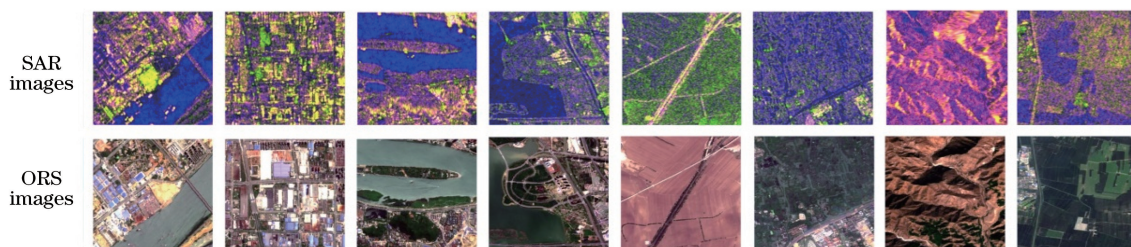


图 11 WHU-SEN-City 数据集图像块样例
Fig. 11 Image blocks of WHU-SEN-City dataset

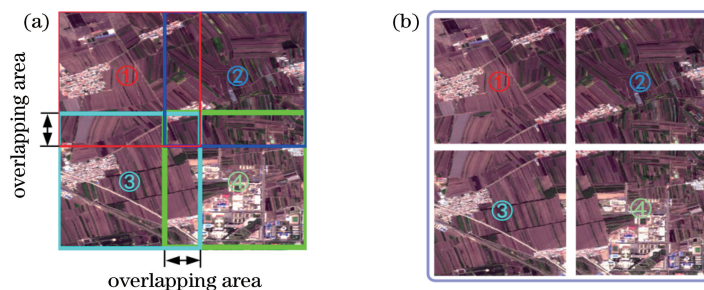


图 12 图像剪切示意图。(a)场景图样例;(b)图像块
Fig. 12 Schematic diagram of image cutting. (a) Example of scene image; (b) image blocks

为了在三个数据集上取得公平的实验数据, 特从各数据集上分别选取 10000 对对应的 SAR 图像与 ORS 图像作为实验对象, 其中 8000 对作为训练集, 剩余 2000 对作为测试集。SEN1-2 数据集的训练集和测试集图像平均取自四个气象季节。SARptical 数据集舍弃 108 对图像, 使用余下的 10000 对图像。WHU-SEN-City 数据集的训练集为来自 26 个城市的 8000 对图像块, 测试集为来自其他 8 个城市的 2000 对图像块。

4.2 网络训练

在模型训练过程中, 网络输入 $256 \text{ pixel} \times 256 \text{ pixel}$ 的图像, 每个数据集上均训练 200 周期, 并使用动量为 0.5 的 Adam 优化器优化训练过程。生成网络和判别网络的学习率均设置为 0.0002, 初始网络参数随机选自均值为 0、方差为 0.02 的高斯分布, 损失函数中的占比参数 λ_{adv} 、 λ_{cyc} 、 λ_{dsc} 、 λ_c 、 λ_s 、 λ_x 、 λ_p 、 λ_{gre} 依次取 1、5、2、2、10、10、10、10。本文所有模型均在配置为 E5-2630v4 CPU、NVIDIA GTX-1080Ti GPU、64 GB RAM 的图形工作站上实现, 以深度学习框架 Tensorflow^[21] 为编译工具。

4.3 评估参数

本文引入峰值信噪比 (PSNR)、结构相似性 (SSIM) 和特征相似性 (FSIM) 三种图像质量评估方法, 对转换模型生成的图像进行定量评价。PSNR 为传统的像质分析指标, 其值越高说明生成图像质量越好。SSIM 和 FSIM 用以计算生成图像与目标

图像之间的相似性, 其值介于 0 到 1 之间。如果生成图像与目标图像相同, 则 SSIM 和 FSIM 等于 1。

5 实验结果与分析

为了验证 SDR-CycleGAN 的有效性, 分别在 SEN1-2、SARptical、WHU-SEN-City 三个数据集上与 MUNIT^[15]、DRIT^[14]、CR-CycleGAN^[22]、S-CycleGAN^[20]、DSMAP^[16] 五种转换模型进行对比实验。

5.1 SEN1-2 数据集实验结果

此实验所选数据样本为 SEN1-2 数据集的一部分, 因此为消除随机性影响, 实验所得的定量比较结果均为在 3 批次不同训练集和测试集上进行实验后的平均结果。

5.1.1 SAR 图像至 ORS 图像转换

不同模型在 SEN1-2 数据集上完成 SAR 图像至 ORS 图像转换的定量比较和对比验证如表 1 和图 13 所示。由表 1 可知, SDR-CycleGAN 生成图像的 PSNR、SSIM 和 FSIM 值均高于其他模型的计算值, 说明所提模型生成的 ORS 图像质量更高, 与目标图像更相似。从图 13 可以看出, DRIT 生成的 ORS 图像色彩单一, 说明其对色彩不敏感。CR-CycleGAN 更加偏重于 SAR 图像着色, 忽略了转换中的结构变换。S-CycleGAN 生成的图像存在较多伪影, 导致图像有所失真。MUNIT 和 DSMAP 生成的图像缺失许多细节信息。而本文模型生成图像的色彩和结构特征与目标 ORS 图像更相近, 且清晰度高、细节完整。

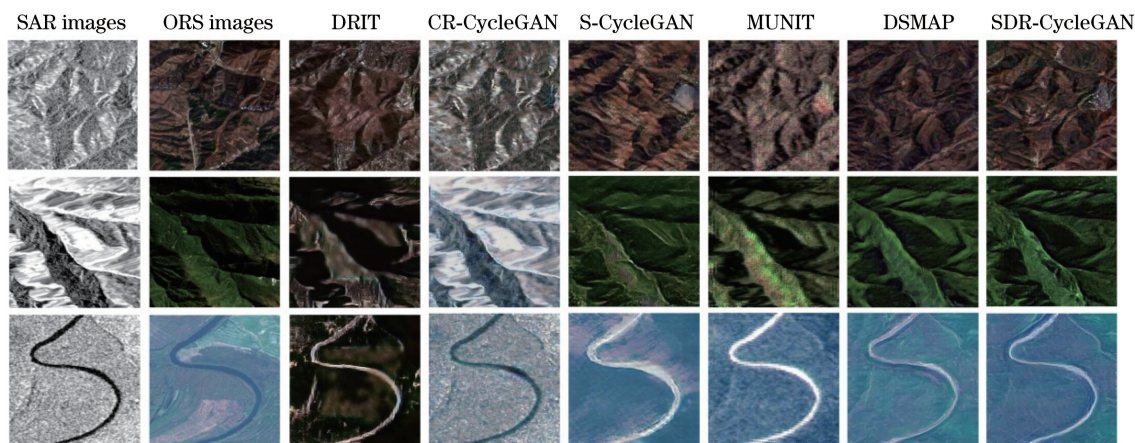


图 13 SEN1-2 数据集 SAR 图像至 ORS 图像转换示例

Fig. 13 Examples of SAR image to ORS image translation on SEN1-2 dataset

5.1.2 ORS 图像至 SAR 图像转换

表 2 比较了各模型在 SEN1-2 数据集上生成 SAR 图像的 PSNR、SSIM 和 FSIM 值, 由其可知,

SDR-CycleGAN 在 ORS 图像至 SAR 图像转换中的表现也优于其他转换模型。图 14 展示了六种模型所生成 SAR 图像的效果对比, 从转换结果来看,

所提模型生成的 SAR 图像未出现失真和伪影现象, 主观判断上最接近于目标 SAR 图像, 但其对山背和道路的刻画有待进一步提高。

表 1 SEN1-2 数据集 SAR 图像至 ORS 图像
转换结果质量评价

Table 1 Quality assessments of SAR image to ORS image translation on SEN1-2 dataset

Method	PSNR /dB	SSIM	FSIM
DRIT	10.140	0.256	0.582
CR-CycleGAN	10.184	0.201	0.629
S-CycleGAN	11.263	0.258	0.664
MUNIT	13.589	0.246	0.670
DSMAP	14.917	0.288	0.653
SDR-CycleGAN	15.466	0.321	0.687

5.2 SARptical 数据集实验结果

对于 SARptical 数据集, 一方面由于 SAR 图像入射角较小且地物多为高层建筑, 致使成像地物大

多呈层状, 另一方面由于 SAR 图像与 ORS 图像成像角度有所偏差, 导致两者共通信息减少, 因此, 使用 SARptical 数据集进行图像转换比 SEN1-2 数据集更具挑战性。此实验选取的样本数量占原数据集样本数量的 98.93%, 其足以代表整个数据集, 因此并未进行重复实验, 实验结果均为一次所得。

表 2 SEN1-2 数据集 ORS 图像至 SAR 图像
转换结果质量评价

Table 2 Quality assessments of ORS image to SAR image translation on SEN1-2 dataset

Method	PSNR /dB	SSIM	FSIM
CR-CycleGAN	9.270	0.231	0.685
S-CycleGAN	10.794	0.261	0.694
DRIT	11.431	0.252	0.682
MUNIT	13.372	0.257	0.689
DSMAP	13.849	0.278	0.701
SDR-CycleGAN	13.912	0.293	0.711

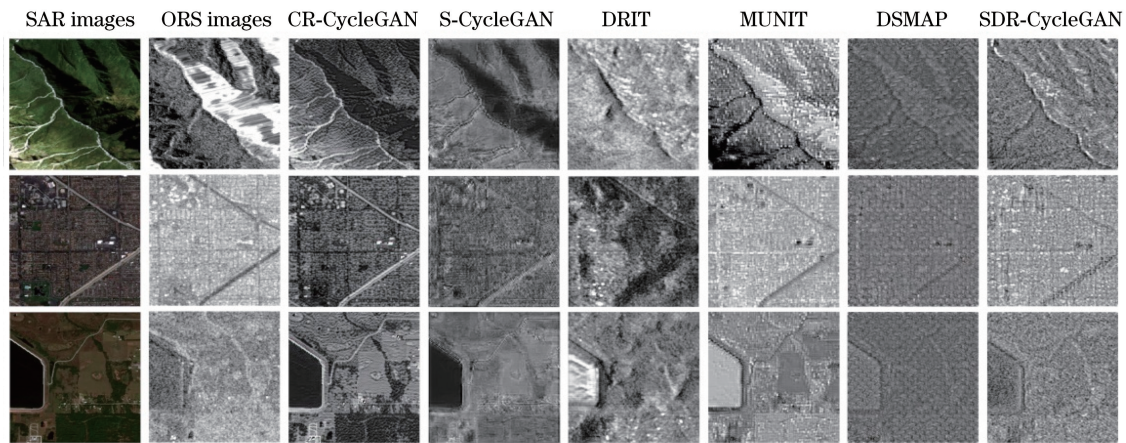


图 14 SEN1-2 数据集 ORS 图像至 SAR 图像转换示例

Fig. 14 Examples of ORS image to SAR image translation on SEN1-2 dataset

5.2.1 SAR 图像至 ORS 图像转换

各模型在 SARptical 数据集上完成 SAR 图像至 ORS 图像转换的量化比较结果见表 3, 转换效果对比见图 15。由表 3 可知, 面对复杂的 SARptical 数据集, SDR-CycleGAN 较其他模型仍表现出明显

表 3 SARptical 数据集 SAR 图像至 ORS 图像
转换结果质量评价

Table 3 Quality assessments of SAR image to ORS image translation on SARptical dataset

Method	PSNR /dB	SSIM	FSIM
DRIT	10.685	0.273	0.652
S-CycleGAN	10.850	0.282	0.658
CR-CycleGAN	11.808	0.311	0.664
MUNIT	11.854	0.287	0.660
DSMAP	12.685	0.289	0.662
SDR-CycleGAN	13.367	0.323	0.687

的优越性。由图 15 可知, DRIT 生成的 ORS 图像色彩单一且地物发生畸变和失真, S-CycleGAN 的转换结果较 DRIT 虽有所改善, 但与理想目标差距较大, CR-CycleGAN 和 MUNIT 的生成图像缺失较多细节特征, 且个别结果存在失真, DSMAP 生成图像的地物表现趋于稳定, 混乱和失真现象得到极大改善, 但与 SDR-CycleGAN 相比, 其对地物信息的刻画仍有所不足。

5.2.2 ORS 图像至 SAR 图像转换

各模型在 SARptical 数据集上 ORS 图像至 SAR 图像转换的定量比较和效果示例, 如表 4 和图 16 所示。由表 4 可知, SDR-CycleGAN 生成图像的 PSNR 和 FSIM 值高于其他模型, 但 SSIM 值低于 CR-CycleGAN。从图 16 可以看出, CR-CycleGAN 生成的 SAR 图像与原始 ORS 图像相比, 只是在图

像色彩上做了转换,结构基本保持一致,地物仍清晰可辨,因此其 SSIM 值在六种模型中最高。但与目标 SAR 图像相比,其生成图像缺失了 SAR 图像特性,FSIM 值在六种模型中最低。SDR-CycleGAN

生成的 SAR 图像不仅地物清晰未失真,且与目标图像特征最为接近,因此判定所提模型与其他模型相比,表现仍最为优异。

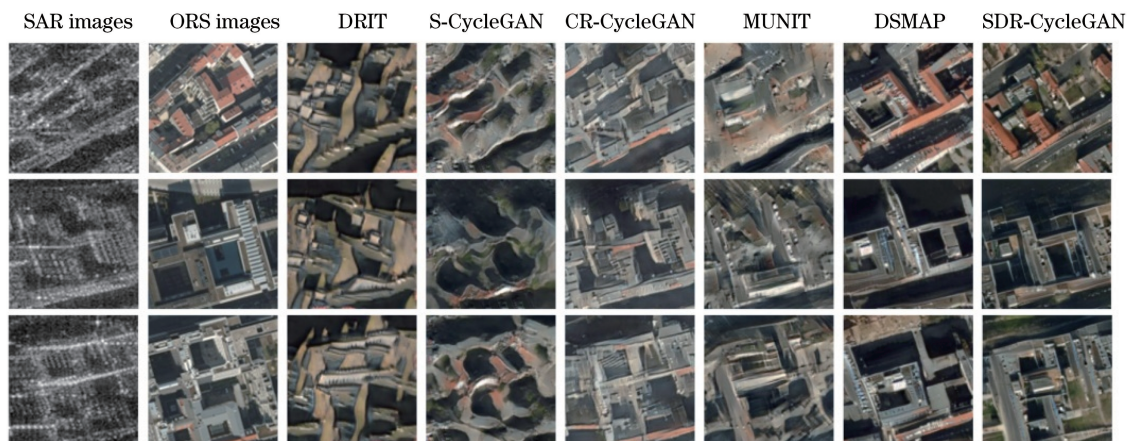


图 15 SARptical 数据集 SAR 图像至 ORS 图像转换示例

Fig. 15 Examples of SAR image to ORS image translation on SARptical dataset

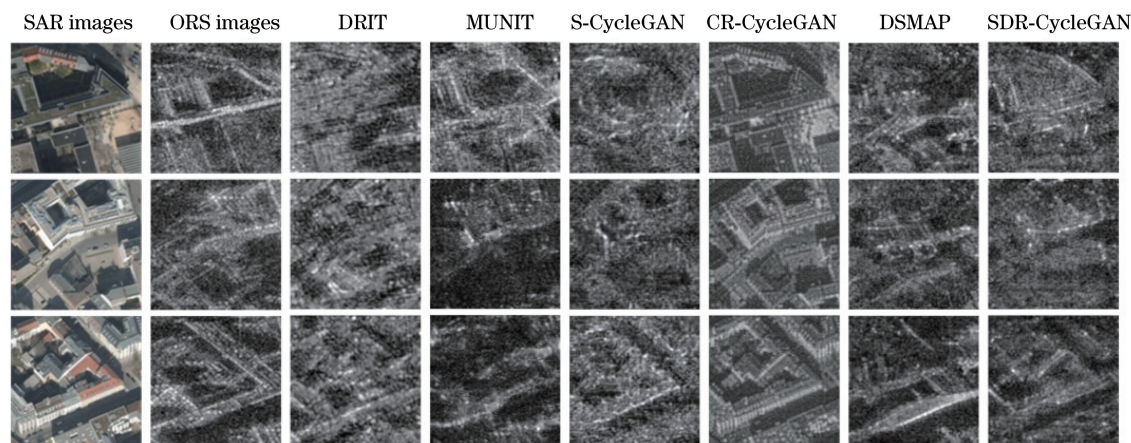


图 16 SARptical 数据集 ORS 图像至 SAR 图像转换示例

Fig. 16 Examples of ORS image to SAR image translation on SARptical dataset

表 4 SARptical 数据集 ORS 图像至 SAR 图像转换结果质量评价

Table 4 Quality assessments of ORS image to SAR image translation on SARptical dataset

Method	PSNR /dB	SSIM	FSIM
DRIT	14.588	0.210	0.723
MUNIT	14.620	0.211	0.720
S-CycleGAN	14.718	0.216	0.725
CR-CycleGAN	15.263	0.257	0.708
DSMAP	15.267	0.240	0.720
SDR-CycleGAN	16.068	0.247	0.728

5.3 WHU-SEN-City 数据集实验结果

SDR-CycleGAN 处理的是 256 pixel × 256 pixel 的图像块,如果要完成大场景的 SAR 图

像与 ORS 图像互转,则需要对转换后的图像进行拼接。但由于生成图像块之间的亮度和色调不同,直接拼接后的场景图会存在明显的接缝线,因此利用加权平均法消除接缝。加权平均法用平均的新像素代替两幅图像的重叠区域,其形式表示为

$$I(x, y) = \{I_1(x, y), (x, y) \in I_1 \xi I_1(x, y) + (1 - \xi)I_2(x, y),$$

$$(x, y) \in I_1 \cap I_2 I_2(x, y), (x, y) \in I_2, (14)$$

式中: $I_1(x, y)$ 和 $I_2(x, y)$ 表示两个相邻图像块 (x, y) 处的像素值; ξ 为控制 I_1 和 I_2 占比的权重,其大小由像素点 (x, y) 距离重叠区域边界的距离决定。

为消除实验数据所产生的随机性影响,此数据集上的定量评估参数值均为在 3 批次不同训练集和

测试集上进行实验后的平均结果。

5.3.1 SAR 图像至 ORS 图像转换

表 5 和图 17 分别列出了六种模型在 WHU-SEN-City 数据集上完成 SAR 图像至 ORS 图像转换的评估参数值和效果示例图。表 5 证明 SDR-CycleGAN 的性能相较于其他模型仍为最优。从图 17 可以看出, BRIT 与 CR-CycleGAN 生成的 ORS 图像在色彩上均与目标 ORS 图像相距甚远, 且 BRIT 生成图像地物混乱失真, CR-CycleGAN 生成图像只完成了简单的色调变换。与前两者相比, MUNIT、S-CycleGAN、DSMAP 和 SDR-CycleGAN 生成的 ORS 图像在整体色彩上更加理想, 但 MUNIT 无法有效区分建筑群、道路、树林、田野等地物目标, S-CycleGAN 对细节地物的刻画存在混

乱和失真现象, DSMAP 对大面积河流色彩转换错误, 且对零星建筑不够敏感, SDR-CycleGAN 生成的 ORS 图像无论是整体还是细节均与目标图像最接近。

表 5 WHU-SEN-City 数据集 SAR 图像至 ORS 图像转换结果质量评价

Table 5 Quality assessment of SAR image to ORS image translation on WHU-SEN-City dataset

Method	PSNR /dB	SSIM	FSIM
DRIT	8.795	0.210	0.565
CR-CycleGAN	11.360	0.248	0.649
MUNIT	13.124	0.317	0.699
S-CycleGAN	13.472	0.391	0.734
DSMAP	13.729	0.447	0.714
SDR-CycleGAN	14.116	0.472	0.763

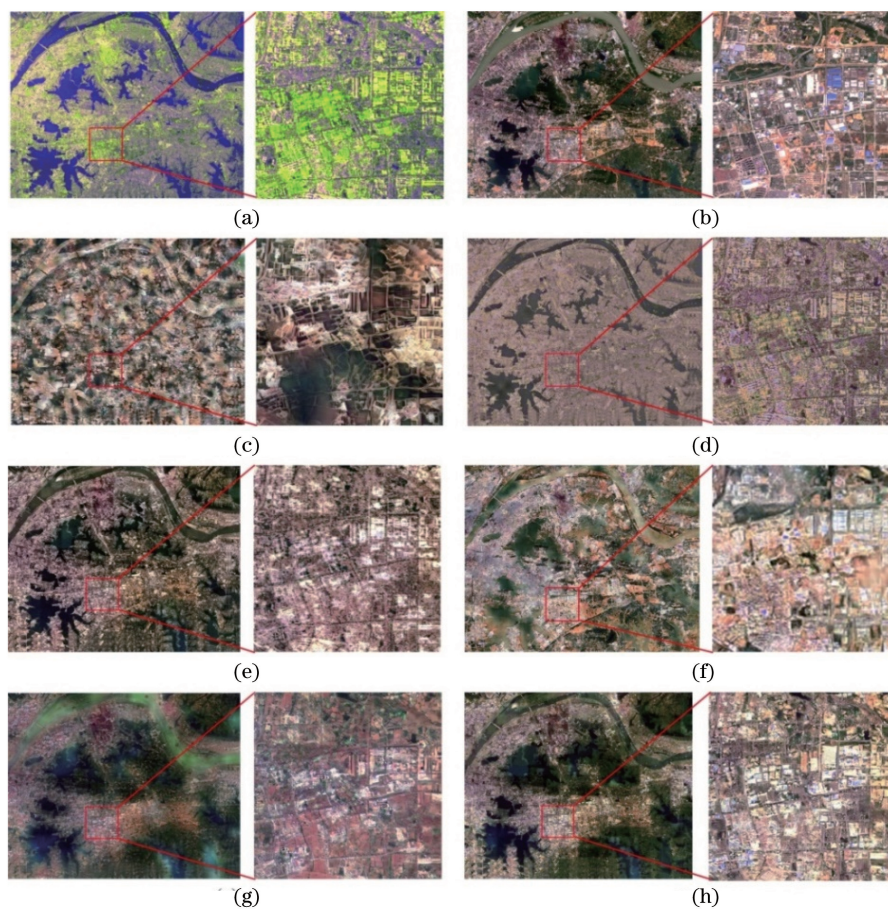


图 17 WHU-SEN-City 数据集 SAR 图像至 ORS 图像转换示例。(a) SAR 图像;(b) ORS 图像;(c) BRIT 转换结果;(d) CR-CycleGAN 转换结果;(e) MUNIT 转换结果;(f) S-CycleGAN 转换结果;(g) DSMAP 转换结果;(h) SDR-CycleGAN 转换结果

Fig. 17 Examples of SAR image to ORS image translation on WHU-SEN-City dataset. (a) SAR images; (b) ORS images; (c) BRIT translation results; (d) CR-CycleGAN translation results; (e) MUNIT translation results; (f) S-CycleGAN translation results; (g) DSMAP translation results; (h) SDR-CycleGAN translation results

5.3.2 ORS 图像至 SAR 图像转换

各模型在 WHU-SEN-City 数据集上完成 ORS

图像至 SAR 图像转换的结果评定和效果展示见表 6 和图 18。由表 6 可知, 相较于其他模型 SDR-

CycleGAN 的表现更加优越。从图 18 可以看出, BRIT 与 CR-CycleGAN 在 ORS 图像至 SAR 图像转换中,仍表现出对色彩不敏感的劣势, BRIT 对建筑物、河流、湖泊、田野等地物的色彩转换与目标图像正相反, CR-CycleGAN 不仅色彩转换错误且地物结构无明显变化。S-CycleGAN 整体转换结果虽与目标接近,但对色调掌控不足,导致“斑块”现象,且地物细节信息转换失败。与前三者相比, MUNIT、DSMAP 和 SDR-CycleGAN 的转换结果较好,但 MUNIT 生成图像模糊,而 DSMAP 转换的地物目标混乱, SDR-CycleGAN 生成的 SAR 图像明显更

加理想。

表 6 WHU-SEN-City 数据集 ORS 图像至 SAR 图像转换结果质量评价

Table 6 Quality assessment of ORS image to SAR image translation on WHU-SEN-City dataset

Method	PSNR /dB	SSIM	FSIM
DRIT	9.076	0.152	0.696
CR-CycleGAN	10.998	0.112	0.687
S-CycleGAN	11.197	0.254	0.682
MUNIT	13.033	0.396	0.711
DSMAP	13.076	0.374	0.697
SDR-CycleGAN	13.170	0.405	0.732

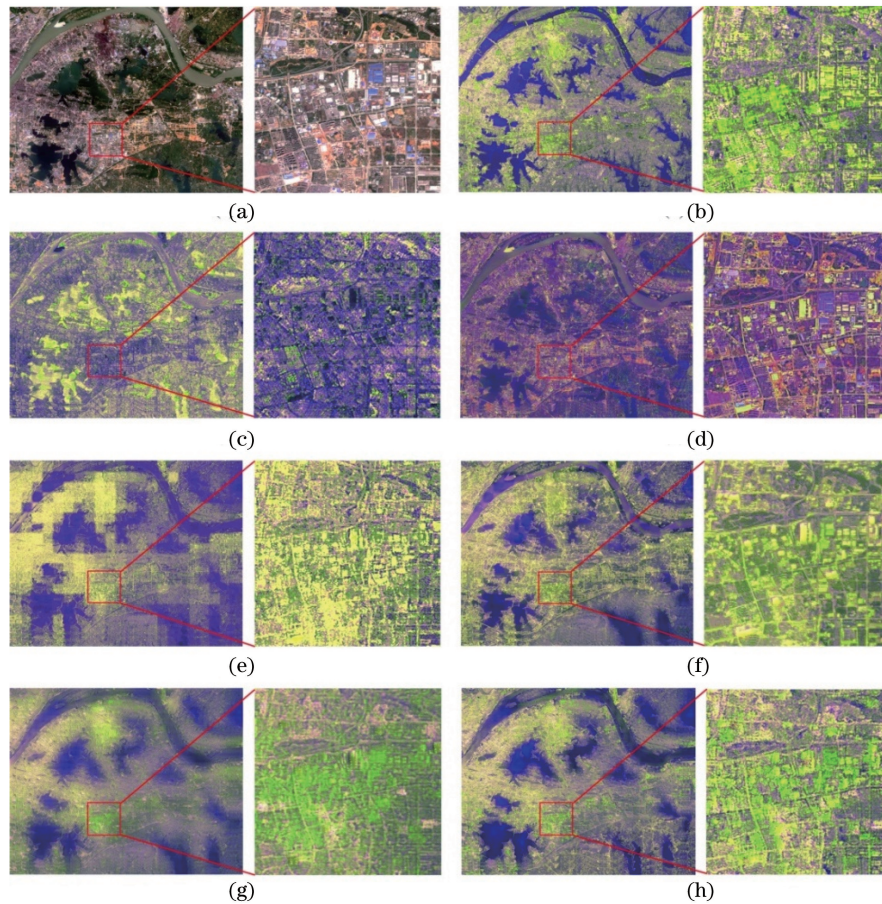


图 18 WHU-SEN-City 数据集 ORS 图像至 SAR 图像转换示例。(a) ORS 图像;(b) SAR 图像;(c) BRIT 转换结果;(d) CR-CycleGAN 转换结果;(e) S-CycleGAN 转换结果;(f) MUNIT 转换结果;(g) DSMAP 转换结果;(h) SDR-CycleGAN 转换结果

Fig. 18 Examples of ORS image to SAR image translation on WHU-SEN-City dataset. (a) SAR images; (b) ORS images; (c) BRIT translation results; (d) CR-CycleGAN translation results; (e) S-CycleGAN translation results; (f) MUNIT translation results; (g) DSMAP translation results; (h) SDR-CycleGAN translation results

通过分析六种图像转换模型在 SEN1-2、SARptical、WHU-SEN-City 三个数据集上完成 SAR 图像与 ORS 图像互转的实验结果可知, DRIT 对图像转换中的色彩变换敏感度低,且生成图像畸变和失真现象严重。CR-CycleGAN 生成的图像色

彩单一,且结构趋向于原始输入图像,其只可完成简单的样式转换,并不能胜任模式差异大的转换任务。S-CycleGAN、MUNIT 和 DSMAP 的性能较前两者有所提高,但对地物细节信息的描述仍有所不足。相较于其他五种模型, SDR-CycleGAN 在任一数据

集上均可出色完成 SAR 图像和 ORS 图像互转,展现出优异的性能和良好的鲁棒性。

6 结 论

本文提出了用于 SAR 图像与 ORS 图像互转的 SDR-CycleGAN 框架,在 DSMAP 基础之上,利用更深的网络层和跳跃连接,分离提取 SAR 图像域和 ORS 图像域的内容特征和风格特征,通过学习两种遥感图像内容特征之间的映射关系,高效完成内容特征转换,而后将输入图像的内容特征与目标域的风格特征相组合,实现图像转换。所提框架还利用 PatchGAN 构建判别器,进一步完善生成图像细节信息。训练过程中,在原网络损失函数基础之上,引入生成重构损失和目标误差损失,减少转换过程中的错误映射和无效信息,使模型更具针对性。通过在 SEN1-2、SARptical、WHU-SEN-City 三个数据集上的对比实验证明,与其他图像转换框架相比,本文方法在主观视觉评价和客观定量分析等方面均取得明显优势,能够生成清晰度更高、结构性更强、特征更加相近的图像。在今后的研究中,将继续探索将该方法应用到更具挑战性的图像转换任务中。

参 考 文 献

- [1] Cheng G, Zhou P C, Han J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(12): 7405-7415.
- [2] Shao J Q, Qu C W, Li J W. A performance analysis of convolutional neural network models in SAR target recognition[C]//2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA), November 13-14, 2017, Beijing, China. New York: IEEE Press, 2017: 1-6.
- [3] Du Z X, Yin J Y, Yang J. Remote sensing aircraft image detection based on semi-supervised learning [J]. Laser & Optoelectronics Progress, 2020, 57(6): 061009.
杜泽星, 殷进勇, 杨建. 基于半监督学习的遥感飞机图像检测方法[J]. 激光与光电子学进展, 2020, 57(6): 061009.
- [4] Yang L, Su J, Huang H, et al. SAR ship detection based on convolutional neural network with deep multiscale feature fusion [J]. Acta Optica Sinica, 2020, 40(2): 0215002.
杨龙, 苏娟, 黄华, 等. 一种基于深层次多尺度特征融合 CNN 的 SAR 图像舰船目标检测算法[J]. 光学学报, 2020, 40(2): 0215002.
- [5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] Fu C Y, Liu W, Ranga A, et al. DSSD: deconvolutional single shot detector[EB/OL]. (2017-01-23) [2020-09-27]. <https://arxiv.org/abs/1701.06659>.
- [7] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [8] Yang X, Sun H, Fu K, et al. Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks [J]. Remote Sensing, 2018, 10(1): 132.
- [9] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C] // The 27th International Conference on Neural Information Processing Systems. Dec. 25-30, Montreal, Canada. Cambridge: MIT, 2014, 2672-2680
- [10] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5967-5976.
- [11] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [12] Hu L M, Zhang Y. Facial image translation in short-wavelength infrared and visible light based on generative adversarial network [J]. Acta Optica Sinica, 2020, 40(5): 0510001.
胡麟苗, 张湧. 基于生成对抗网络的短波红外-可见光人脸图像翻译[J]. 光学学报, 2020, 40(5): 0510001.
- [13] Yi Z L, Zhang H, Tan P, et al. DualGAN: unsupervised dual learning for image-to-image translation[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2868-2876.
- [14] Lee H Y, Tseng H Y, Huang J B, et al. Diverse

- image-to-image translation via disentangled representations[C]//Computer Vision - ECCV 2018, 2018.
- [15] Huang X, Liu M Y, Belongie S, et al. Multimodal unsupervised image-to-image translation[M]. Cham: Springer International Publishing, 2018: 179-196.
- [16] Chang H Y, Wang Z X, Chuang Y Y. Domain-specific mappings for generative adversarial style transfer [M]. Cham: Springer International Publishing, 2020: 573-589.
- [17] Huang X, Belongie S. Arbitrary style transfer in real-time with adaptive instance normalization [C]// 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1510-1519.
- [18] Schmitt M, Hughes L H, Zhu X X. The SEN1-2 dataset for deep learning in SAR-optical data fusion [J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018: 141-146.
- [19] Wang Y Y, Zhu X X. The SARoptical dataset for joint analysis of SAR and optical image in dense urban area [C] // IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, July 22-27, 2018, Valencia, Spain. New York: IEEE Press, 2018: 6840-6843.
- [20] Wang L, Xu X, Yu Y, et al. SAR-to-optical image translation using supervised cycle-consistent adversarial networks [J]. IEEE Access, 2019, 7: 129136-129149.
- [21] Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems [EB/OL]. (2016-03-16) [2020-09-18]. <https://arxiv.org/abs/1603.04467>.
- [22] Fu S L, Xu F, Jin Y Q. Reciprocal translation between SAR and optical remote sensing images with cascaded-residual adversarial networks [EB/OL]. (2019-10-09) [2020-09-18]. <https://arxiv.org/abs/1901.08236>.