

引用格式: LI Bing, XIAN Yong, ZHANG Daqiao. Infrared Image Generation Algorithm Based on Conditional Generation Adversarial Networks[J]. Acta Photonica Sinica, 2021, 50(11):1110004

李冰, 鲜勇, 张大巧. 基于条件生成对抗网络的红外图像生成算法[J]. 光子学报, 2021, 50(11):1110004

基于条件生成对抗网络的红外图像生成算法

李冰, 鲜勇, 张大巧

(火箭军工程大学 作战保障学院, 西安 710025)

摘 要: 针对实地采集红外数据成本高、难度大, 且实效性较低的问题, 提出一种基于条件生成对抗网络的红外图像生成算法。在条件生成对抗网络架构的基础上, 利用 D-LinkNet 构建生成网络, 从而较好地学习图像的纹理和图像间的依赖关系; 通过 PatchGAN 构建判别网络, 进而有效地处理图像的高频分量, 并降低计算量。同时, 结合批归一化处理技术优化训练过程, 缓解生成对抗网络模式坍塌现象。最后在制作的红外/可见光数据集上进行实验验证, 实验结果证明, 本文算法能够生成高质量可靠的红外基准图。

关键词: 红外图像生成; 生成对抗网络; 生成模型; 判别模型; 图像质量评估

中图分类号: TP18

文献标识码: A

doi: 10.3788/gzxb20215011.1110004

Infrared Image Generation Algorithm Based on Conditional Generation Adversarial Networks

LI Bing, XIAN Yong, ZHANG Daqiao

(College of War Support, Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: Collecting infrared images from fields is difficult, high-costing and time-consuming. In order to address this problem, an infrared image generation method based on conditional generative adversarial networks is proposed. In the proposed method, the D-LinkNet network is utilized to build the generative model, enabling improved learning of rich image textures and identification of dependencies between images. Moreover, the PatchGAN architecture is employed to build a discriminant model to process the high-frequency components of the images effectively and reduce the amount of calculation required. In addition, batch normalization is used to optimize the training process, and thereby the instability and mode collapse of the generated adversarial network training can be alleviated. Finally, experimental verification is conducted on the produced infrared/visible light dataset. The experimental results reveal that high-quality and reliable infrared data are generated by the proposed algorithm.

Key words: Infrared image generation; Generative adversarial networks; Generative model; Adversarial model; Image quality assessment

OCIS Codes: 100.3008; 110.3080; 150.1135; 350.2660; 110.3000

0 引言

受应用背景和保障能力的限制, 红外成像制导中的基准图通常是可见光图像, 而实时图为红外图像。

基金项目: 国家自然科学基金(Nos. 41574008, 61302195, 41774156)

第一作者: 李冰(1993—), 女, 博士研究生, 主要研究方向为人工智能在图像生成、目标检测与识别方面的应用。Email: libingbenyi@163.com

导 师: 鲜勇(1972—), 男, 教授, 博士, 主要研究方向为制导、突防与任务规划、弹道设计以及人工智能在目标识别与导航方面的应用。

Email: xy603xy@163.com

收稿日期: 2021-03-09; 录用日期: 2021-07-13

<http://www.photon.ac.cn>

红外图像与可见光图像的成像机理不同,导致两者间存在较大的特征差异,进而增加了红外成像制导中景象匹配的难度。利用红外图像仿真技术生成所需环境下场景的红外特性,不仅能够有效地降低获取红外数据的成本,还可以生成很多自然环境以及场景条件下外场试验难以获得的红外数据。生成的红外数据能够应用于航空、航海、气象、地质以及农业等多个领域,为探测、分类、定位、识别、跟踪等任务提供基础、可靠的数据。

近年来,随着计算机性能的不断提高以及深度学习理论的快速发展和完善,涌现出许多新的基于神经网络的生成模型,其中生成对抗网络(Generative Adversarial Networks, GANs)^[1]能够满足许多领域的研究和应用需求,已经成为人工智能领域中最重要研究热点之一。ANTIPOV G等^[2]利用条件生成对抗网络(Conditional Generative Adversarial Networks, CGAN)生成人脸图像。双路径生成对抗网络(Two-Pathway Generative Gdversarial Networks, TP-GAN)^[3]与表示解析生成对抗网络(Disentangled Representation Learning-Generative Adversarial Networks, DR-GAN)^[4]将GANs应用于人脸转正领域(人脸转正是从单一侧脸图像合成未高清的正面人脸图像的技术)。根据马尔科夫原理的马尔可夫生成对抗网络(Markovian Generative adversarial networks, MGAN)^[5]在生成图像纹理方面与Texture Network^[6]具有相同的合成速度。ISOLA P等提出Pix2pix实现了黑白图到彩色图、卫星图到地图、语义图到街景图、边缘图到照片的转换^[7]。BROCK A等用BigGAN^[8]生成的图片纹理和背景比较逼真,但算法复杂性较高。图像超分辨率生成对抗网络(Image Super Resolution Generative adversarial networks, SRGAN)利用残差网络和VGG(Visual Geometry Group, VGG)网络分别作为生成器和判别器得到了较好的纹理细节学习效果^[9]。文献[10]中利用对偶学习的DualGAN模型实现了跨域图像生成。在对偶学习的基础上,KARRAS T等将循环一致性引入到生成对抗网络中,扩展了图像到图像的转换工作^[11]。文献[12]在给来自Kinect的RGB图像与曲线法线贴图的情况下,以表面法线作为生成对抗网络的条件来生成图像,该模型被称作Style-GAN。基于Style-GAN模型,YANG和LIM提出了一种框架,该框架能够生成与给定的一次性照片示例相同分布的面部图像^[13]。

目前,大部分研究是将GANs用于人脸合成、纹理生成、可见光到夜视图像等,而国内外将GANs模型用于红外图像仿真领域的研究比较缺乏。鉴于外场试验获得红外数据成本太高、难度相对较大、数据相对较少的情况,本文提出一种基于CGAN模型的红外图像生成算法,用于在可见光图像基础上仿真生成红外图像,来提供可靠的红外数据、扩大红外数据库、制备红外基准图。在CGAN框架的基础上,D-LinkNet网络来构建生成网络,输入为可见光图像,输出为红外仿真样本;再将真实目标样本与生成的仿真样本用于训练基于PatchGAN的判别网络,输出样本属于相应类别的概率。通过交替迭代训练生成网络和判别网络,最终生成的红外仿真样本与真实样本的数据分布基本相同。

1 算法原理

GANs由GOODFELLOW I J等在2014年的国际神经信息处理系统大会上提出,它是在深度生成模型的基础上发展而来的一种新的生成模型。GANs与其他生成模型的主要区别在于:GANs使用对抗的方式,先通过判别器来学习生成样本与训练样本间的差异,再引导生成器去减小这种差异;其他生成模型则主要是直接以数据分布和模型分布的差异为目标函数。

1.1 条件生成对抗网络

GANs的本质是:通过两个玩家间的对抗来实现学习,完成零和博弈。在零和博弈中,一名玩家作为生成器,其主要任务是生成与真实样本数据分布尽可能一致的数据,进而蒙骗另一位玩家。另外一名玩家作为判别器,其目的是准确地判断输入的样本是否属于真实样本数据。通常,学者们将生成网络与对抗网络比喻成假钞制造者与验钞机。GANs的训练过程类似于假钞制造者与验钞机之间的博弈:假钞制造者尽可能地提高假钞的制作水平,生成与真钞尽可能相同的假钞,来躲过验钞机的检验;而验钞机则不断地提高辨别假钞的能力。随着GANs的不断训练,假钞制造者制作假钞的能力与验钞机的辨别能力都在不断提高^[14]。

GANs中包含两个网络:生成网络(生成器 G)和对抗网络(判别器 D),对应的是生成模型和对抗模型。

GANs的结构如图1。

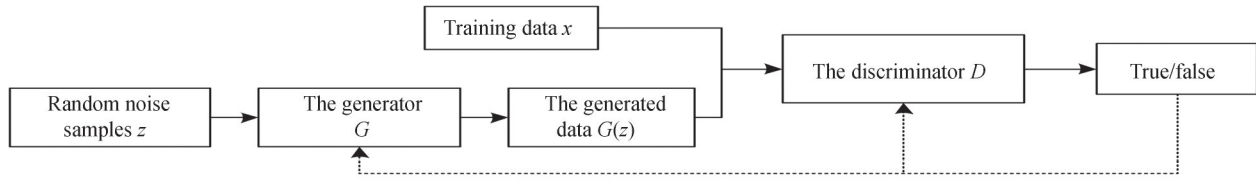


图1 GANs的基本结构框架

Fig. 1 The basic framework of the original GANs

在原始GANs中,GOODFELLOW IJ定义GANs的值函数^[1]为

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

式中, $x \sim p_{\text{data}}$ 表示 x 取自真实分布的数据, $z \sim p_z$ 表示随机噪声 z 取自模拟分布的数据(例如高斯噪声分布), (\cdot) 表示期望值。

针对原始GANs不能生成具有特定属性的图片的问题, MIRZA M等^[15]提出了条件生成对抗网络(Conditional Generative Adversarial Networks, CGAN)。CGAN的核心在于将条件信息 y 融入到生成器和判别器中, 条件 y 可以是任何标签信息, 比如人脸图像的面部表情、图像类别等。CGAN网络结构如图2。

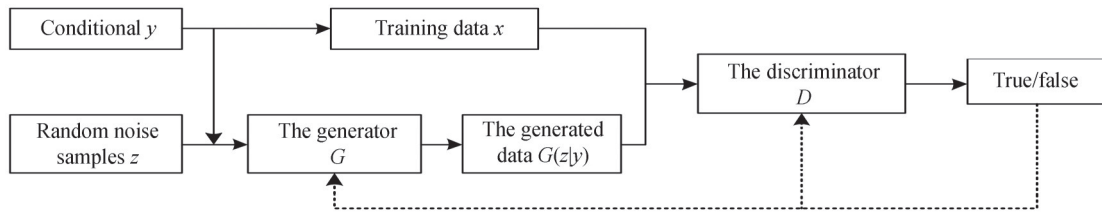


图2 CGAN的基本结构

Fig. 2 The basic framework of CGAN

1.2 损失函数

CGAN的损失函数为

$$\ell_{\text{CGAN}}(G, D) = E_{x,y} [\log D(x, y)] + E_{x,z} \{\log \{1 - D[x, G(x, z)]\}\} \quad (2)$$

在CGAN网络框架基础上, 提出的算法用图像代替随机噪声作为输入。为了更好地利用输入图像的结构信息, 在损失函数中引入了L1损失函数。

$$\ell_{L1}(G) = E_{x,y,z} \left[\left\| y - G(x, z) \right\|_1 \right] \quad (3)$$

本文算法的损失函数最终定义为

$$G^* = \arg \min_G \max_D \ell_{\text{CGAN}}(G, D) + \lambda E_{x,y,z} \left[\left\| y - G(x, z) \right\|_1 \right] \quad (4)$$

1.3 基于D-Linknet的生成网络

与常见的先下采样到低维, 再上采样到原始分辨率的编码-解码结构的网络相比, 以LinkNet为基本框架, 并引入了空洞卷积的D-LinkNet^[16]具有跳跃链接(用来保留不同分辨率下像素级的细节信息)、残差块和编码-解码体系结构的优势, 能够增加网络的接收域, 保留图像的空间细节信息, 并实现多尺度特征融合。

在本文算法中, 使用D-LinkNet构造生成网络, 并且D-LinkNet被设计为接收 256×256 大小的图片作为输入。如图3, D-LinkNet由A、B、C三部分组成, 其分别是编码部分、中心部分和解码部分。编码部分使用在ImageNet数据集上训练好的ResNet34作为网络的编码器, 解码部分使用转置卷积进行上采样。在中心部分增加带有shortcut的空洞卷积, 这能增强网络识别能力、扩大接受域、融合多尺度信息。中心部分的展开结构如图4, 从上到下空洞卷积层的空洞数分别是2、1和0, 其对应的感受野分别是7、3和1。

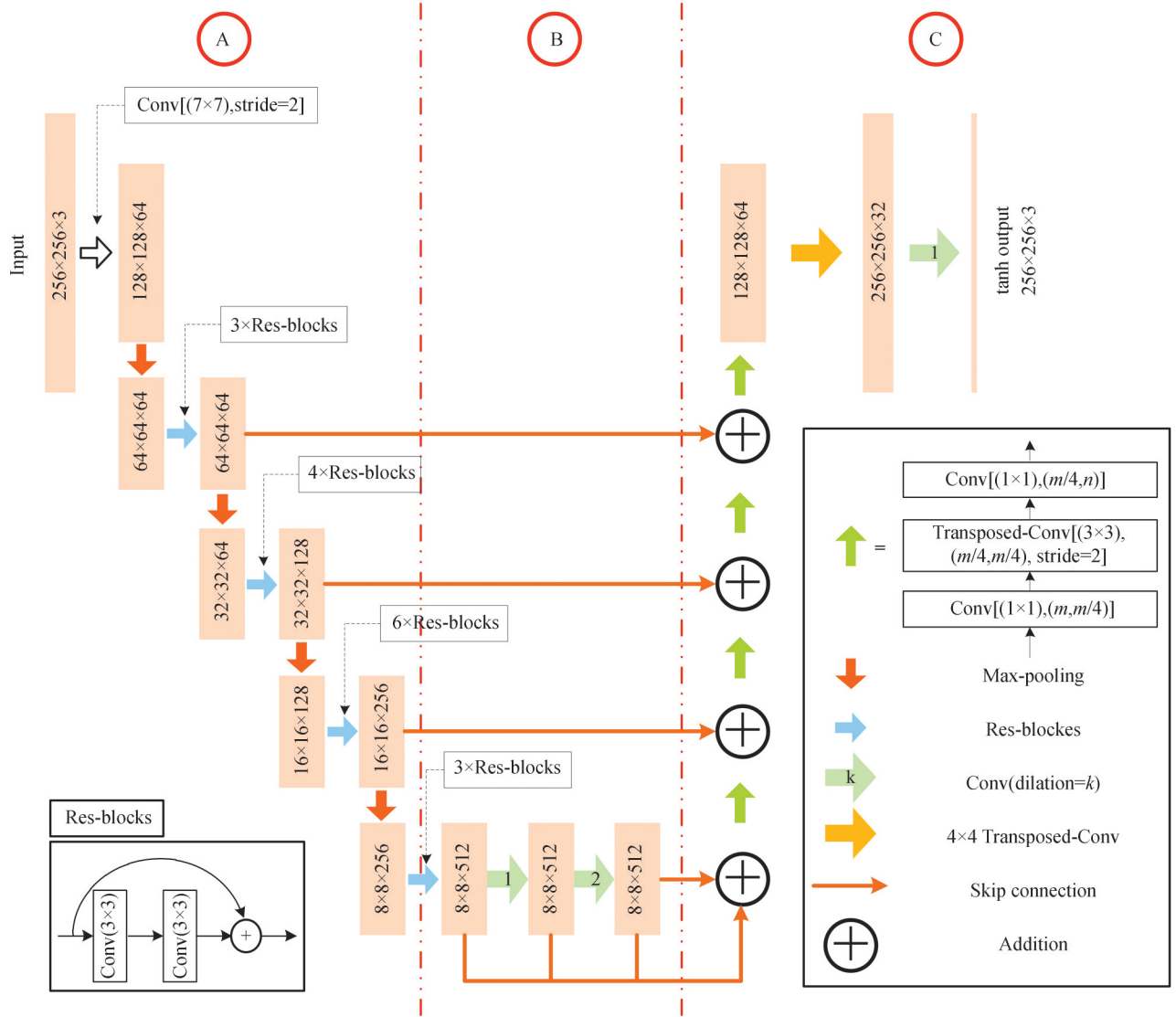


图3 D-LinkNet 结构

Fig. 3 Structure of D-LinkNet

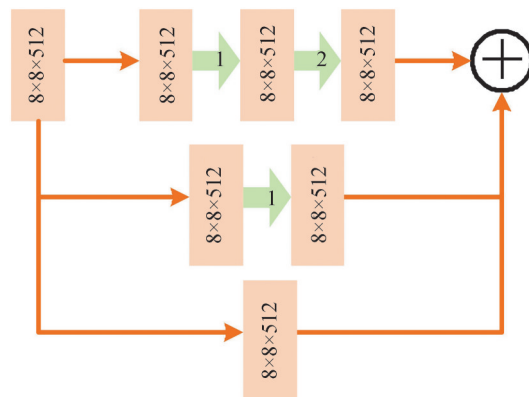


图4 D-LinkNet 中心部分的展开结构

Fig. 4 The unrolled structure of the center dilation part of D-LinkNet

1.4 基于 PatchGAN 的判别网络

在本算法中,使用卷积“PatchGAN”(Markovian discriminator)分类器构建判别网络。马尔科夫判别器

(Markovian discriminator)是一种判别模型,它由卷积层组成,输出是 $N \times N$ 的矩阵 X ,最后取 X 矩阵的平均值作为真/假的输出。因为输出矩阵中的每一个值对应着原图中的一个感受野,即对应了原图中的一个区域块(patch),因而具备该结构的GANs亦被称为PatchGAN。

以输入大小为 256×256 的图像为例,常规的GANs是将大小的图像映射到单个标量输出,表示“真”或者“假”,但这不易体现图像的局部特征;而PatchGAN则是将 256×256 大小的图像映射到大小为 $N \times N$ 的矩阵,其中每个 X_{ij} 表示图像中的区域块 ij 是真或假。PatchGAN实现了图像局部特征的提取及表征,使模型能更加关注图像的细节信息,有利于生成更高质量的图像。

由于引入L1项作为辅助来学习低频信息,能够获得低频信息的准确率,所以本文算法的判别器只需要对图像的高频成分进行建模构建高频信息。因此,本文算法中,采用PatchGAN构造判别器。在算法结构中,PatchGAN判别器的输入是多个固定大小的图像块。该判别器关注的是图像的局部结构,只需要判断图像中每个区域块的真假即可,其仅在图像块的尺度上惩罚结构,能够有效减小训练中的参数数量。

1.5 算法训练过程

算法1:生成对抗网络的小批量随机梯度下降训练。其中,步骤数 k 是超参数,本文算法的实验中设置 $k=1$ 。

for 训练迭代次数 do

for k 步 do

从 $p_G(z)$ 数据分布中采样 m 个样本 $\{z^{(1)}, \dots, z^{(m)}\}$;

从 $p_{\text{data}}(x)$ 数据分布中采样 m 个样本 $\{x^{(1)}, \dots, x^{(m)}\}$;

通过增加随机梯度来更新判别器:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left\{ \log D(x^{(i)}) + \log \left\{ 1 - D[G(z^{(i)})] \right\} \right\}$$

end for

从 $p_G(z)$ 分布中采样 m 个样本 $\{z^{(1)}, \dots, z^{(m)}\}$;

通过降低随机梯度来更新生成器:

$$\nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m \log \left\{ 1 - D[G(z^{(i)})] \right\}$$

end for

基于梯度的更新可以使用任何标准的基于梯度的学习规则。本文算法的实验中使用了动量。

2 实验验证与分析

2.1 数据集介绍

利用无人机搭载同轴红外成像仪和可见光成像仪对指定区域进行同时拍摄,采集红外/可见光数据。红外/可见光数据采集试验工作选用大疆经纬 M200 V2 无人机作为搭载平台,选用 DJI Zenmuse XT2 热成像云台相机(该云台相机上同轴搭载了红外成像仪和可见光相机,能同时拍摄红外图像和可见光图像。)采集数据时选择的气象条件包含晴天、多云、阴天、雾霾、雨天和雪天;数据采集的时间横跨春、夏、秋、冬四个季节;拍摄时间包含早上、中午和傍晚;拍摄视角至少包括3个,并且同一拍摄视角上至少采集两种不同距离上的图像数据。目标包括建筑物(材质包括钢质、混凝土、水泥、砖混)、雷达天线(罩)、冷却塔、烟囱、风力发电机、油库、跑道、船只等。数据中的场景包含城市、校区、街道、工厂、居民区、交通枢纽、河面等。

从泛化能力方面出发,多采用随机选取的方法选取数据。为了验证本文算法的有效性,在上述红外/可见光数据集中随机选取1374组红外/可见光数据进行实验,并按照1070:304的比例构建训练样本集和测试样本集。选取的1374组数据中,目标包括建筑物(材质包括钢质、混凝土、水泥、砖混)、雷达天线(罩)、冷却塔、烟囱、风力发电机、油库、跑道、船只等,选取的数据中涵盖的气象条件包含晴天、多云、阴天、雾霾、雨、雪。

2.2 实验验证与分析

为了验证算法的有效性,利用采集的数据进行了大量实验,并从目视判读分析和基于质量指标的客观验证两方面对红外图像的生成效果进行评价。

2.2.1 目视判读分析

基于本文算法的建筑物、烟囱、冷却塔的红外生成图像实例依次如图5~7所示,其中图5中的建筑物材质包括钢质、混凝土、水泥和砖混。根据红外图像的先验知识和目视判读可以看出:

- 1) 算法能够较好地生成晴天、薄雾、阴天、多云、霾等气象干扰条件下,目标、场景的红外效果。
- 2) 算法能够有效生成不同材质、不同外观形状的建筑物的红外图像,如图5所示。图中各类材质的建筑物的红外生成图像细节突出、轮廓可以辨别、灰度信息与真实的红外效果比较接近。

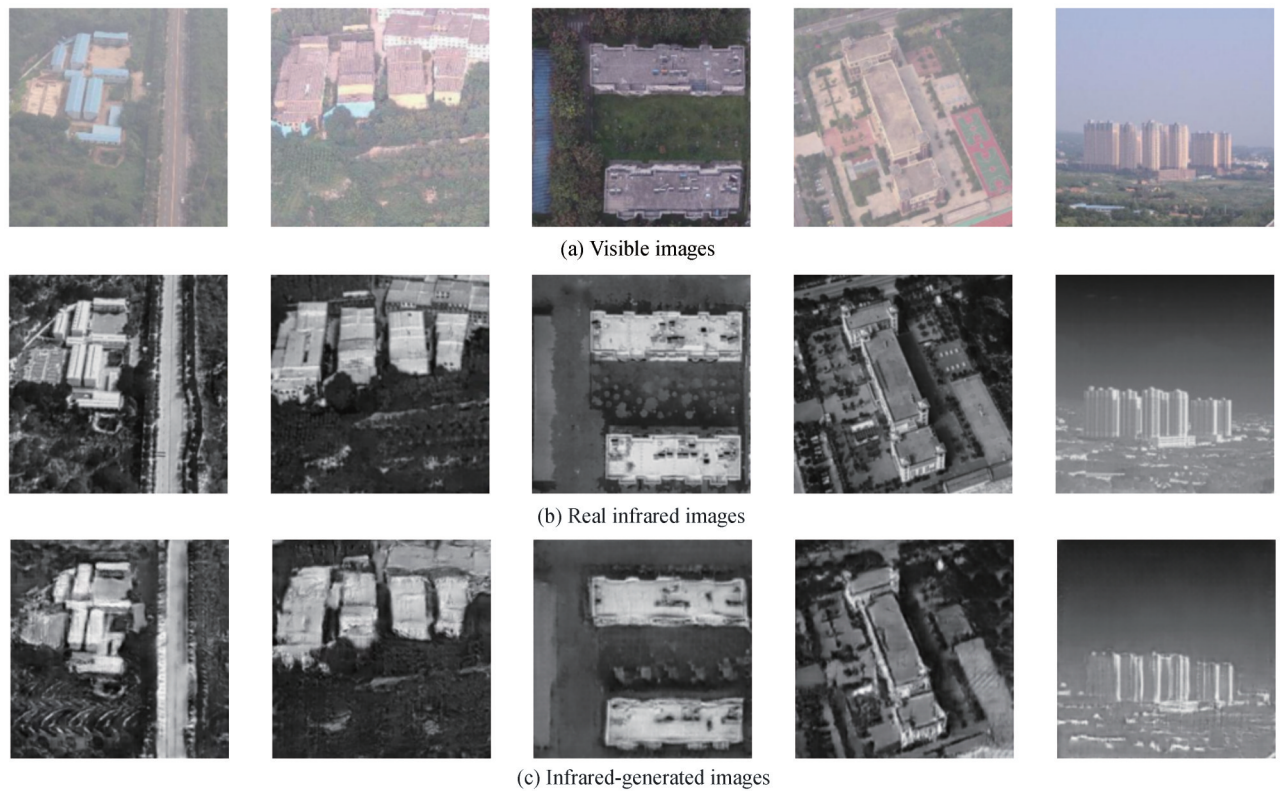
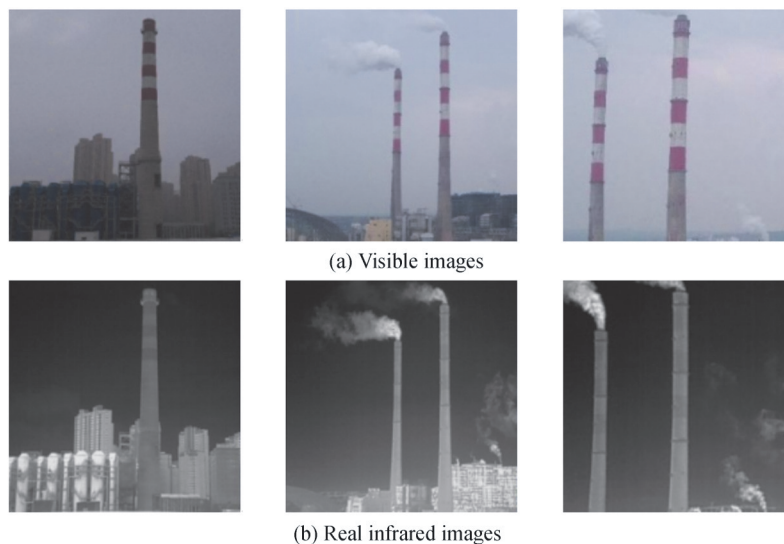


图5 基于本文算法的建筑物红外生成实例

Fig. 5 Examples of infrared-generated images of buildings produced on the basis of our method



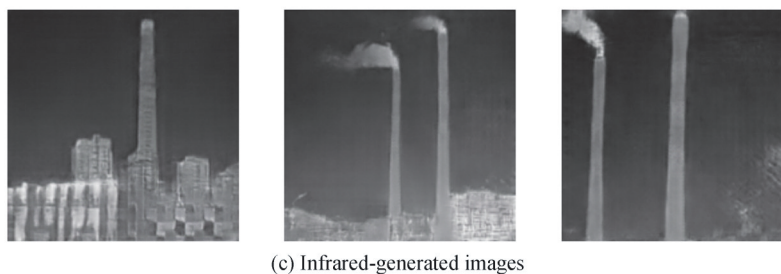


图6 基于本文算法的烟囱红外生成实例

Fig. 6 Examples of infrared-generated images of chimneys produced on the basis of our method

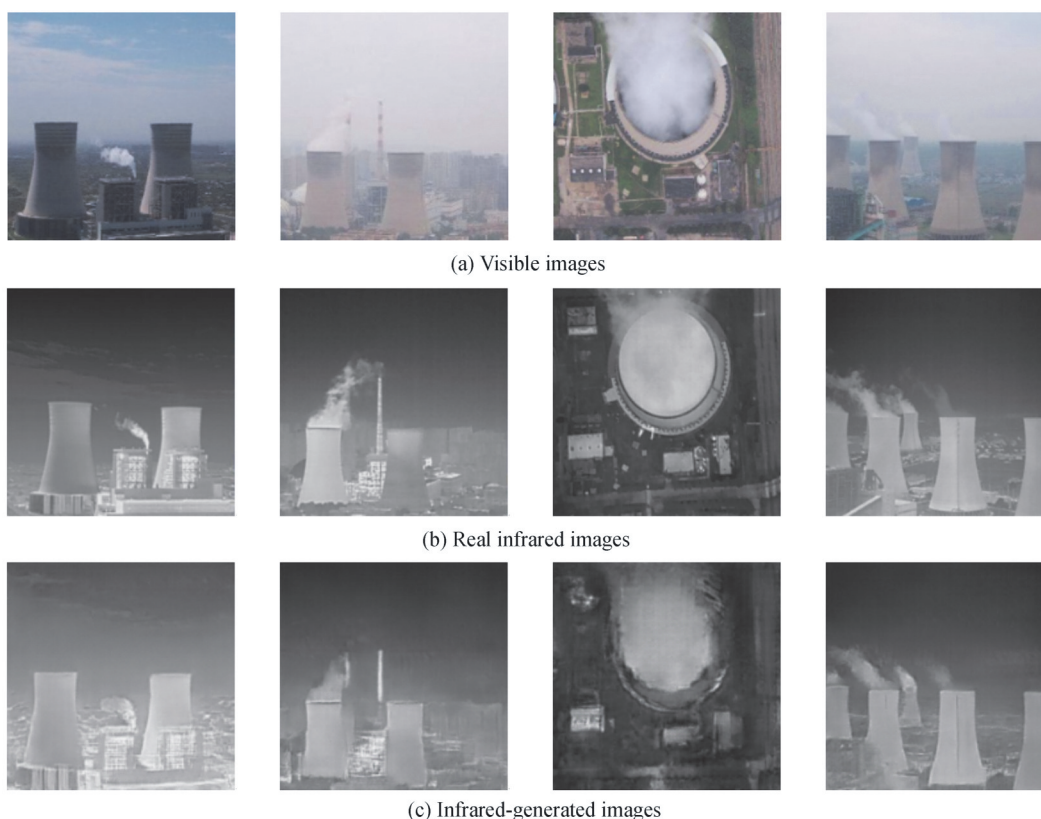


图7 基于本文算法的冷却塔红外生成实例

Fig. 7 Examples of infrared-generated images of cooling towers produced on the basis of our method

3)算法能够很好地生成烟雾(热气)的红外特性,如图6和图7所示。根据红外生成图中烟囱或冷却塔的顶部是否存在烟雾的红外特性,可以辨别出烟囱或冷却塔是否处于工作状态。

通过目视判读以及主观评价可知:本文算法生成的红外图像具有较好的纹理信息,并且其灰度信息和轮廓信息与真实的红外图像较接近,两者的相似度较高,红外生成效果较好。

2.2.2 基于质量指标的图像评估分析

一般情况下,红外生成图与红外实时图之间灰度特性的相似性越大,说明红外图像的生成效果越好。为了客观地评价本文算法生成红外图像的效果,通过计算红外生成基准图(红外生成基准图是从红外生成结果图中通过人机交互分割出来的)和红外实时图之间的均方根误差(Radical Mean Square Error, RMSE)和特征相似性(Feature SIMilarity, FSIM)^[17]来评价红外图像的生成效果。

均方根误差是两幅图像信息变化程度的测量,该值能反映出两幅图之间的灰度值差异。通常,均方根误差值越小,表示两者灰度差异越小,即反映出红外图像的仿真效果越好;反之,则反映出红外图像的仿真效果越差。特征相似性是对结构相似性的改进,既利用相位一致性提取丰富的纹理、边缘和结构信息,又引入梯度幅度提取图像的对比信息,能较好地评价图像间的结构差异。一般情况下,特征相似性的值越大,表

示图像间相似性越大,即反映出红外仿真效果越好。因为用户更注重目标的红外生成效果,因此只计算目标的红外实时图与红外生成图之间的均方根误差和特征相似性。均方根误差和特征相似性分别按照式(5)、式(6)计算,其中 I 和 S 分别表示目标的红外实测图和红外仿真图, $PC_1(I)$ 和 $PC_2(S)$ 分别表示 I 和 S 的相位一致性, $G_1(I)$ 和 $G_2(S)$ 分别表示 I 和 S 的梯度幅度。

$$RMSE = \left(\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (I(x,y) - S(x,y))^2 \right)^{1/2} \quad (5)$$

$$\begin{cases} S_{PC} = \frac{2PC_1(I) \cdot PC_2(S) + T_1}{PC_1^2(I) + PC_2^2(S) + T_1} \\ S_G = \frac{2G_1(I) \cdot G_2(S) + T_2}{G_1^2(I) + G_2^2(S) + T_2} \\ S_L = [S_{PC}]^\alpha \cdot [S_G]^\beta \\ FSIM = \frac{\sum S_L \cdot \max(PC_1(I), PC_2(S))}{\sum \max(PC_1(I), PC_2(S))} \end{cases} \quad (6)$$

为验证红外图像的生成效果,分别与以U-net256、ResNet9、ResNet34模块为生成器的生成对抗网络算法进行实验对比。其中以U-net256模块为生成器的生成对抗网络即Pix2pix算法^[7],用“Pix2pix”来描述;以ResNet9、ResNet34模块为生成器的生成对抗网络算法分别用“Resnet9”和“Resnet34”来描述。四种对比算法的网络结构如表1。

表1 四种GANs算法的网络组成
Table 1 The network structure of the four GANs algorithms

Method	Networks				
	Generator				Discriminator
	U-net256	ResNet9	ResNet34	D-LinkNet	PatchGAN
Pix2pix	✓				✓
Resnet9		✓			✓
Resnet34			✓		✓
Our method				✓	✓

数据集中共有1374组红外/可见光图像,按照1070:304的比例构建训练样本集和测试样本集。均方根误差指标中,指标值小的一方为优;特征相似性指标中,指标值大的一方为优。对图像质量评价指标的实际值进行优劣数量统计,定义统计结果为优劣比。本文统计了这四种算法生成的所有红外图像与对应的真实红外图像之间的均方根误差以及特征相似性指标值,并计算各指标值的平均值(分别用mRMSE和mFSIM表示)以及算法间图像质量指标值的优劣比,统计结果如表2。

依据表2给出的实验数据可以得出:

- 1) 本文算法结果的均方根误差平均值(mRMSE)为33.82,在四种算法中最小,指标评价相对最优;
- 2) 本文算法结果的特征相似性平均值(mFSIM)为0.737,在四种算法中最大,指标评价相对最优;

表2 图像质量评价指标的平均值和优/劣样本数
Table 2 The average and the superior/inferior sample numbers of the evaluation indexes

Method	The average of the evaluation indexes		The superior/inferior sample numbers of the evaluation indexes			
	mRMSE	mFSIM	RMSE Samples		FSIM Samples	
			Our' < Other'	Our' > Other'	Our' > Other'	Our' < Other'
Pix2pix	35.01	0.721	207	97	220	84
Resnet9	34.42	0.722	180	124	220	84
Resnet34	37.04	0.700	228	76	243	61
Our method	33.82	0.737	\	\	\	\

3) 在 304 组对比数据中,本文算法的 RMSE 指标值中优于 Pix2pix 算法的样本数量为 207 个,优于 Resnet9 算法的样本数量为 180 个,优于 Resnet34 算法的样本数量为 228 个;

4) 在 304 组对比数据中,本文算法的 FSIM 指标值中优于 Pix2pix 算法的样本数量为 220 个,优于 Resnet9 算法的样本数量为 220 个,优于 Resnet34 算法的样本数量为 243 个;

根据上述数据分析可得,本文算法生成的红外图像的质量优于其他三种生成对抗网络算法。

1) 均方根误差值统计结果

为了更加直观地表示实验结果,以本文算法得出的 304 个均方根误差值大小顺序为基准,绘制出了本文算法与 Pix2pix 算法实验结果的对比图,本文算法的实验结果用曲线“—★—”表示, Pix2pix 算法的实验结果用散点“■”表示。如图 8。

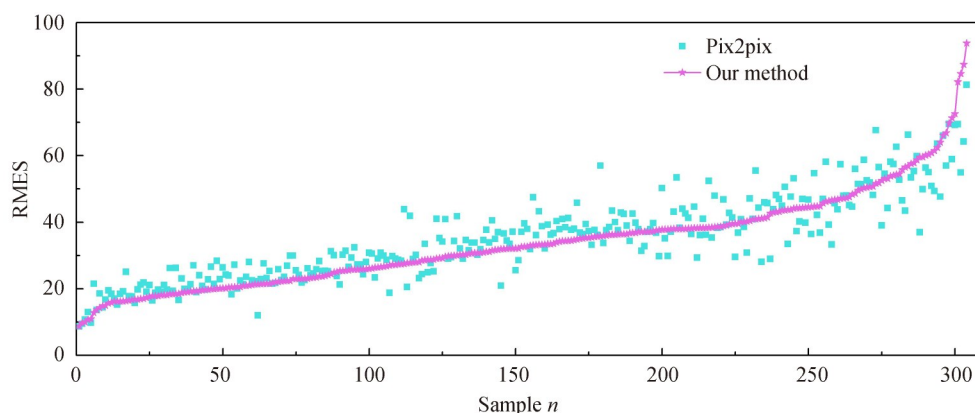


图 8 本文算法与 Pix2pix 算法生成红外图像与真实红外图像间的均方根误差结果

Fig. 8 The RMSE results between the real infrared images and the infrared images generated by our method and the Pix2pix

由图 8 可得,分布在曲线“—★—”以上部分的散点“■”,明显多于曲线以下部分。本文算法的均方根误差指标结果中,有 207 个指标值优于 Pix2pix 算法、97 个指标值劣于 Pix2pix 算法,即两种算法的均方根误差指标优劣比为 207:97,说明本文算法生成的红外图像中有 207 张图像的质量优于 Pix2pix 算法。

本文算法、Resnet9 以及 Resnet34 算法得出的均方根误差指标结果如图 9。图中本文算法、Resnet9 和 Resnet34 的均方根误差值分别用曲线“—★—”、散点“●”、散点“▲”表示。

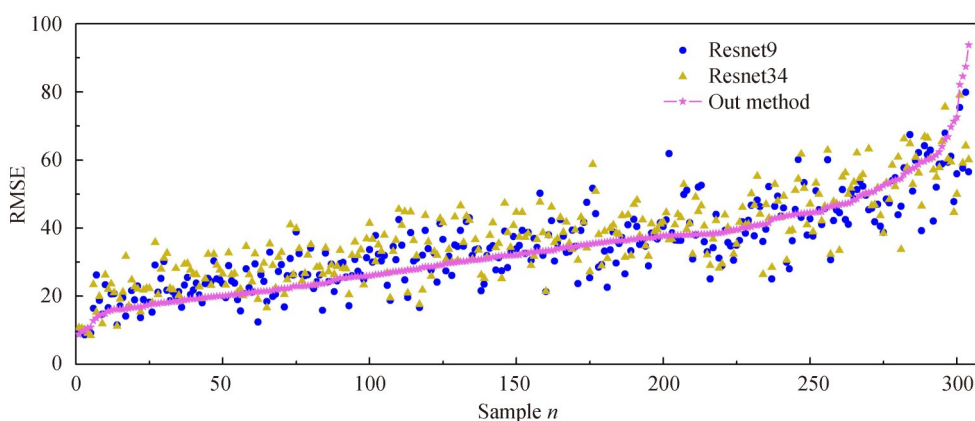


图 9 本文算法与 Resnet9、Resnet34 算法生成红外图像与实际红外图像之间灰度值的均方根误差结果

Fig. 9 The RMSE results between the real infrared images and the infrared images generated by our method and the Resnet9 and Resnet34

由图 9 可得,分布在曲线“—★—”以上部分的散点“●”和“▲”,明显多于曲线以下部分。本文算法与 Resnet9 算法的均方根误差指标优劣比为 180:124,与 Resnet34 算法的优劣比为 228:76,即说明本文算法生成的红外图像的质量明显优于 Resnet9、Resnet34 算法。

2) 特征相似性统计结果

本文算法与另外三种算法的特征相似性指标结果如图10。图中本文算法、Pix2pix算法、Resnet9算法和Resnet34算法的特征相似性值分别用曲线“—◆—”、散点“■”、散点“●”和散点“▲”表示。

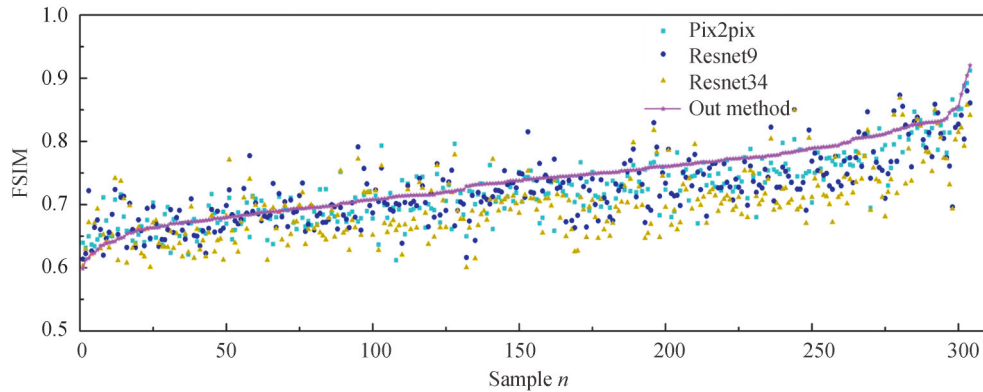


图10 本文算法与另外三种算法生成红外图像与实际红外图像之间特征相似性结果

Fig. 10 The FSIM results between the real infrared images and the infrared images generated by our method and the other three methods

由图可得,分布在曲线“—◆—”以下部分的“■”、“●”和“▲”散点,明显多于曲线以上部分。本文算法的特征相似性指标结果中,有220个指标值优于Pix2pix算法、84个指标值劣于Pix2pix算法,即两种算法的特征相似性指标优劣比为220:84,说明本文算法生成的红外图像中有220张图像的质量优于Pix2pix算法。同理本文算法与Resnet9算法的特征相似性指标优劣比为220:84,与Resnet34算法的优劣比为243:61,即说明本文算法生成的红外图像的质量明显优于Resnet9算法和Resnet34算法。

均方根误差需要考虑两幅图像对应点的灰度值。由于可见光图像与红外实测图像之间存在尺度变换、旋转、方位角等差异,虽然经过了配准,但是仍然存在部分配准误差,导致不能实现完全的精确配对,从而影响了均方根误差的计算,可能导致均方根误差值偏大;特征相似性考虑的是两幅图像的全局信息的分布情况,受图像间配准误差的影响较小。由图8至图10可知,相比于其它三种GANs算法,本文算法在整体上均方根误差值偏小、特征相似值偏大;并且RMSE与FSIM两个指标的优劣比最大值分别为3和3.98,均远大于1,则可说明本文算法的生成效果优于Pix2pix算法以及分别以ResNet9和ResNet34模块为生成器的生成对抗网络。综上,根据目视判读分析和基于质量指标的图像评估分析可知:本文算法生成的红外数据与红外实测图之间相似度较高,能够较好地反映出目标的红外特性,算法的红外生成效果较好。

3 结论

本文提出了一种基于条件生成对抗网络的红外图像生成算法。该算法引入D-LinkNet网络构建生成网络来学习图像纹理并发现图像间的依赖关系,采用PatchGAN构造判别模型,以便于有效地处理图像的高频成分、降低计算量;在训练过程中,结合批归一化和Adam优化训练过程,来缓解训练不稳定以及模式坍塌现象。在制作的红外/可见光图像数据集上进行仿真验证得出,本文算法生成的红外图像较逼真,能够较好地反映出目标的红外特性,与真实的红外图像相似度较高,算法的红外图像生成效果较好。

参考文献

- [1] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]. Proceedings of the 2014 International Conference on Neural Information Processing Systems (NeurIPS), MIT Press, 2014: 2672-2680.
- [2] ANTIPOC G, BACCOUCHE M, DUGELAY J L. Face aging with conditional generative adversarial networks[C]. Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), 2017: 2089-2093.
- [3] HUANG Rui, ZHANG Shu, LI Tianyu, et al. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis[C]. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, 2017: 2458-2467.
- [4] TRAN L, YIN Xi, LIU Xiaoming. Disentangled representation learning gan for pose-invariant face recognition[C].

- Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017: 1283-1292.
- [5] LI Chuan, WANG M. Precomputed real-time texture synthesis with markovian generative adversarial networks [C]. Proceedings of the 2016 European Conference on Computer Vision (ECCV), IEEE, 2016: 702-716.
- [6] ULAANOV D, LEBEDEV V, VEDALDI A, et al. Texture networks: feed-forward synthesis of textures and stylized images[C]. Proceedings of the 2016 International Conference on Machine Learning (ICML), 2016: 1349-1357.
- [7] ISOLA P, ZHU Junyan, ZHOU Tinghui, et al. Image-to-image translation with conditional adversarial networks [C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017: 5967-5976.
- [8] BROCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis[J]. arXiv preprint arXiv:1809.11096.
- [9] LDEIG C, THEIS L, HUSZSR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017: 105-114.
- [10] YI Zili, ZHANG Hao, TAN Ping, et al. DualGAN: unsupervised dual learning for image-to-image translation [C]. Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, 2017: 2868-2876.
- [11] ZHU Junyan, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. Proceedings of the 2017 IEEE Conference on Computer Vision (ICCV), IEEE, 2017: 2223-2232.
- [12] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks [C]. Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2019: 4401-4410.
- [13] YANG Chao, LIM Sernam. One-Shot domain adaptation for face generation [C]. Proceedings of the 2020 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2020: 5920-5929.
- [14] GOODFELLOW I. NIPS 2016 tutorial: Generative adversarial networks [C]. Proceedings of the 2016 International Conference on Neural Information Processing Systems (NeurIPS), arXiv preprint arXiv:1701.00160.
- [15] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. arXiv preprint arXiv:2014.1411.1784.
- [16] ZHOU Lichen, ZHANG Chuang, WU Ming. D-LinkNet: linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, doi:10.1109/cvprw.2018.00034.
- [17] ZHANG Lin, ZHANG Lei, MOU Xuanqi, et al. FSIM: a feature similarity index for image quality assessment [J]. IEEE Transactions on Image Processing, 2011, 20(8): 2378-2386.