



JOURNAL OF IMAGE AND GRAPHICS

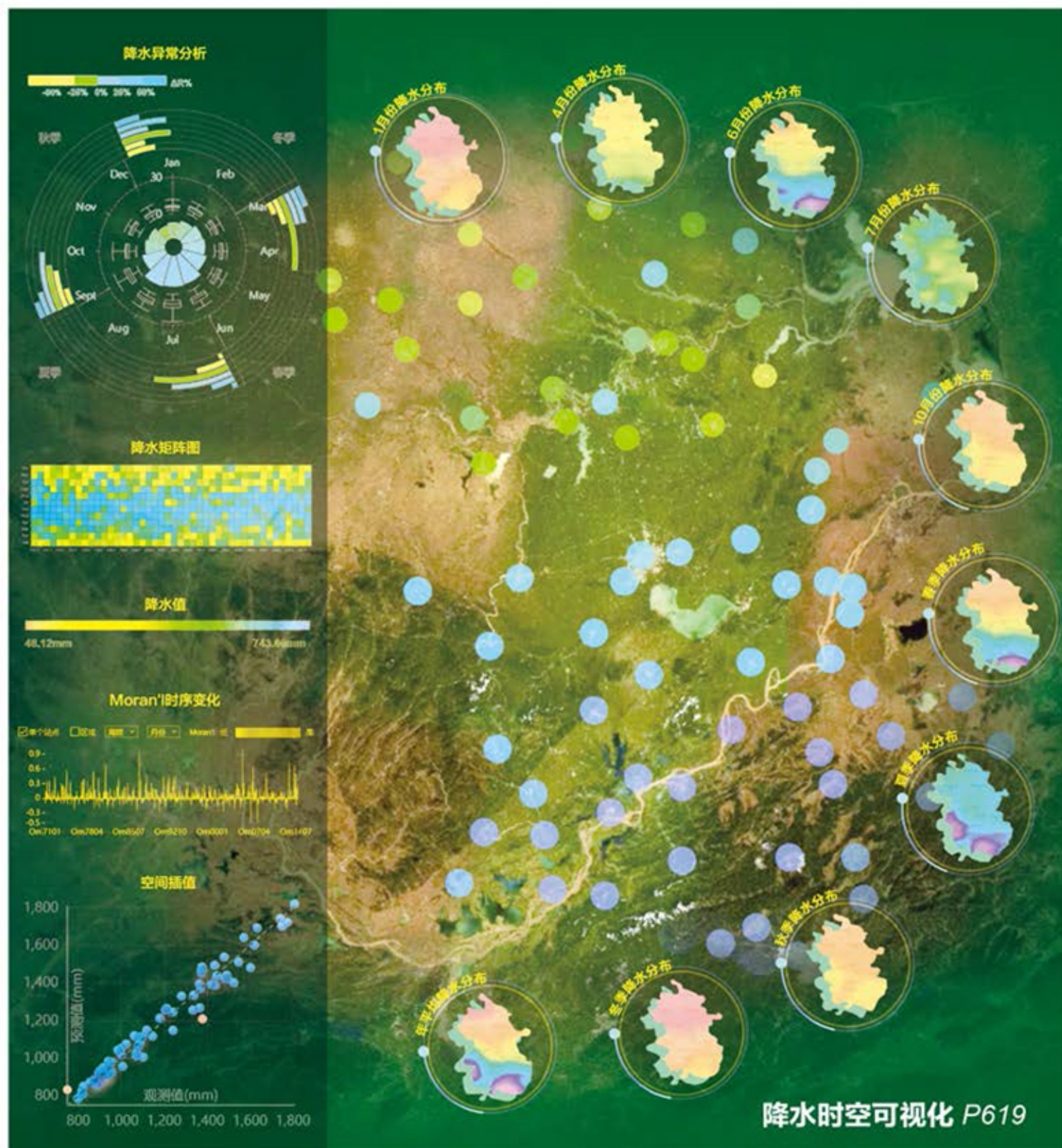
主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报

图形学

2021
03
VOL.26

ISSN1006-8961
CN11-3758/TB



中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）



第26卷第3期（总第299期）

2021年3月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑部
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@radi.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

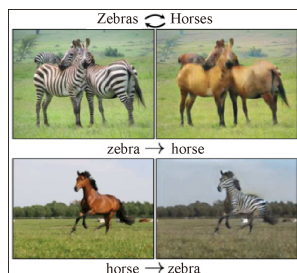
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial Office of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

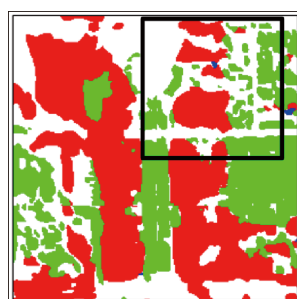
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



深度学习中的图像数据增广方法研究(第0487页)



多通道递归残差网络的图像超分辨重建(第0605页)



面向GF-2遥感影像的U-Net城市绿地分类(第0700页)

学者观点

深度学习与生物医学图像分析2020年综述

陈弘扬, 高敬阳, 赵地, 汪红志, 宋红, 苏庆华 0475

综述

深度学习中的图像数据增广方法研究综述

马崇鼻, 唐娉, 赵理君, 张正 0487

图像处理和编码

自适应B样条小波函数模糊距离测量

梁锐, 魏阳杰 0503

图像分析和识别

增强二阶网络调制的目标跟踪

王献海, 宋慧慧, 张开华, 刘青山 0516

背景与方向感知的相关滤波跟踪

姜文涛, 涂湖, 刘万军 0527

SSD与时空特征融合的视频目标检测

尉婉青, 禹晶, 柏鳗晏, 肖创柏 0542

自然场景下变形车牌检测模型DLPD-Net

余焯, 付源梓, 陈维笑, 刘海涛 0556

两阶段特征提取策略的图像去雾

袁非牛, 李志强, 史劲亭, 夏雪, 李雅 0568

图像理解和计算机视觉

由粗到精的多尺度散焦模糊检测

衡红军, 叶何斌, 周末, 黄睿 0581

Re-GAN: 残差生成式对抗网络算法

史彩娟, 涂冬景, 刘靖祯 0594

多通道递归残差网络的图像超分辨重建

程德强, 郭昕, 陈亮亮, 寇旗旗, 赵凯, 高蕊 0605

计算机图形学

降水时空关联特征可视分析

周志光, 谢琬滢, 郑微桦, 陈圆圆 0619

医学图像处理

结合残差路径及密集连接的乳腺超声肿瘤分割

陈杨怀, 陈胜, 姚莉萍 0633

遥感图像处理

增强小目标特征的航空遥感目标检测

赵文清, 孔子旭, 周震东, 赵振兵 0644

密集子区域切割的任意方向舰船快速检测

陈华杰, 吴栋, 谷雨 0654

CGAN样本生成的遥感图像飞机识别

王耀领, 王宏琦, 许滔 0663

高分辨率遥感影像的边缘损失增强地物分割

陈琴, 朱磊, 吕燧栋, 吴谨 0674

含多级通道注意力机制的CGAN遥感图像建筑物分割

余帅, 汪西莉 0686

面向GF-2遥感影像的U-Net城市绿地分类

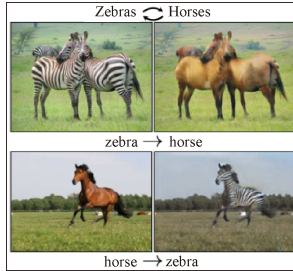
徐知宇, 周艺, 王世新, 王丽涛, 王振庆 0700

条件生成对抗遥感图像时空融合

李昌洁, 宋慧慧, 张开华, 张晓露, 刘青山 0714

CONTENTS

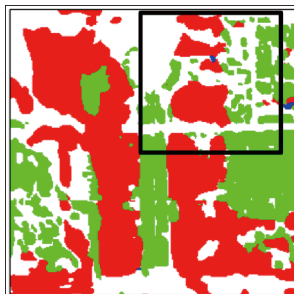
JOURNAL OF IMAGE AND GRAPHICS



Review of data augmentation for image in deep learning(P0487)



Image super-resolution reconstruction from multi-channel recursive residual network (P0605)



U-Net for urban green space classification in Gaofen-2 remote sensing images(P0700)

Scholar View

Review of the research progress in deep learning and biomedical image analysis till 2020
Chen Hongyang, Gao Jingyang, Zhao Di, Wang Hongzhi, Song Hong, Su Qinghua 0475

Review

Review of data augmentation for image in deep learning
Ma Dongao, Tang Ping, Zhao Lijun, Zhang Zheng 0487

Image Processing and Coding

Adaptive distance measurement method with blur of B-spline wavelet function
Liang Rui, Wei Yangjie 0503

Image Analysis and Recognition

Object tracking using enhanced second-order network modulation
Wang Xianhai, Song Huihui, Zhang Kaihua, Liu Qingshan 0516

Background and direction-aware correlation filter tracking
Jiang Wentao, Tu Chao, Liu Wanjun 0527

Video object detection using fusion of SSD and spatiotemporal features
Yu Wanqing, Yu Jing, Bai Manyan, Xiao Chuangbai 0542

DLPD-Net: distorted license plate detection model in natural scenarios
Yu Ye, Fu Yuanzi, Chen Weixiao, Liu Haitao 0556

Image defogging algorithm using a two-phase feature extraction strategy
Yuan Feiniu, Li Zhiqiang, Shi Jinting, Xia Xue, Li Ya 0568

Image Understanding and Computer Vision

Coarse-to-fine multiscale defocus blur detection
Heng Hongjun, Ye Hebin, Zhou Mo, Huang Rui 0581

Re-GAN: residual generative adversarial network algorithm
Shi Caijuan, Tu Dongjing, Liu Jingyi 0594

Image super-resolution reconstruction from multi-channel recursive residual network
Cheng Deqiang, Guo Xin, Chen Liangliang, Kou Qiqi, Zhao Kai, Gao Rui 0605

Computer Graphics

Visual analytics of spatial and temporal correlation features of rainfall
Zhou Zhiguang, Xie Wanying, Zheng Weihua, Chen Yuanyuan 0619

Medical Image Processing

Tumor segmentation in breast ultrasound combined with Res paths and a dense connection
Chen Yanghuai, Chen Sheng, Yao Liping 0633

Remote Sensing Image Processing

Target detection algorithm of aerial remote sensing based on feature enhancement technology
Zhao Wenqing, Kong Zixu, Zhou Zhendong, Zhao Zhenbing 0644

Fast detection algorithm for ship in arbitrary direction with dense subregion cutting
Chen Huajie, Wu Dong, Gu Yu 0654

Aircraft recognition of remote sensing image based on sample generated by CGAN
Wang Yaoling, Wang Hongqi, Xu Tao 0663

Segmentation of high-resolution remote sensing image by collaborating with edge loss enhancement
Chen Qin, Zhu Lei, Lyu Suidong, Wu Jin 0674

Remote sensing building segmentation by CGAN with multilevel channel attention mechanism
Yu Shuai, Wang Xili 0686

U-Net for urban green space classification in Gaofen-2 remote sensing images
Xu Zhiyu, Zhou Yi, Wang Shixin, Wang Litao, Wang Zhenqing 0700

Spatiotemporal fusion of satellite images via conditional generative adversarial learning
Li Changjie, Song Huihui, Zhang Kaihua, Zhang Xiaolu, Liu Qingshan 0714

中图法分类号: TP753 文献标识码: A 文章编号: 1006-8961(2021)03-0663-11

论文引用格式: Wang Y L, Wang H Q and Xu T. 2021. Aircraft recognition of remote sensing image based on sample generated by CGAN. Journal of Image and Graphics, 26(03):0663-0673 (王耀领, 王宏琦, 许滔. 2021. CGAN 样本生成的遥感图像飞机识别. 中国图象图形学报, 26(03):0663-0673) [DOI:10.11834/jig.200001]

CGAN 样本生成的遥感图像飞机识别

王耀领^{1,2,3,4}, 王宏琦^{1,4}, 许滔^{1,2,3,4}

1. 中国科学院空天信息创新研究院, 北京 100190; 2. 中国科学院大学电子电气与通信工程学院, 北京 100049;
3. 中国科学院大学, 北京 100049; 4. 中国科学院空天信息创新研究院网络信息体系技术重点实验室, 北京 100190

摘要: 目的 基于深度学习的飞机目标识别方法在遥感图像解译领域取得了很大进步,但其泛化能力依赖于大规模数据集。条件生成对抗网络(conditional generative adversarial network, CGAN)可用于产生逼真的生成样本以扩充真实数据集,但对复杂遥感场景的建模能力有限,生成样本质量低。针对这些问题,提出了一种结合 CGAN 样本生成的飞机识别框架。**方法** 改进条件生成对抗网络,利用感知损失提高生成器对遥感图像的建模能力,提出了基于掩膜的结构相似性(structural similarity, SSIM)度量损失函数(masked-SSIM loss)以提高生成样本中飞机区域的图像质量,该损失函数与飞机的掩膜相结合以保证只作用于图像中的飞机区域而不影响背景区域。选取一个基于残差网络的识别模型,与改进后的生成模型结合,构成飞机识别框架,训练过程中利用生成样本代替真实的卫星图像,降低了对实际卫星数据规模的需求。**结果** 采用生成样本与真实样本训练的识别模型在真实样本上的进行实验,前者的准确率比后者低 0.33%;对于生成模型,在加入感知损失后,生成样本的峰值信噪比(peak signal to noise ratio, PSNR)提高了 0.79 dB, SSIM 提高了 0.094;在加入基于掩膜的结构相似性度量损失函数后,生成样本的 PSNR 提高了 0.09 dB, SSIM 提高了 0.252。**结论** 本文提出的基于样本生成的飞机识别框架生成了质量更高的样本,这些样本可以替代真实样本对识别模型进行训练,有效地解决了飞机识别任务中的样本不足问题。

关键词: 深度学习;卷积神经网络;生成对抗网络;光学遥感图像;目标识别

Aircraft recognition of remote sensing image based on sample generated by CGAN

Wang Yaoling^{1,2,3,4}, Wang Hongqi^{1,4}, Xu Tao^{1,2,3,4}

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China; 4. Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China

Abstract: **Objective** Aircraft type recognition is a fundamental problem in remote sensing image interpretation, which aims to identify the type of aircraft in an image. Aircraft type recognition algorithms have been widely studied and improved ceaselessly. The traditional recognition algorithms are efficient, but their accuracy is limited by the small capacity and poor robustness. The deep-learning-based methods have been widely implemented because of good robustness and generalization,

收稿日期:2020-01-08;修回日期:2020-06-09;预印本日期:2020-06-15

基金项目:国家自然科学基金项目(41701508)

Supported by: National Natural Science Foundation of China (41701508)

especially in the object recognition task. In remote sensing scenes, the objects are sparsely distributed; hence, the available samples are few. In addition, labeling is time consuming, resulting in a modest number of labeled samples. Generally, the deep-learning-based models rely on a large amount of labeled data due to thousands of weights needed to learn. Consequently, these models suffer from scarce data that are insufficient to meet the demand of large-scale datasets, especially in the remote sensing scene. Generative adversarial network (GAN) can produce realistic synthetic data and enlarge the scale of the real dataset. However, these algorithms usually take random noises as input; therefore, they are unable to control the position, angle, size, and category of objects in synthetic images. Conditional GAN (CGAN) have been proposed by previous researchers to generate synthetic images with designated content in a controlled scheme. CGANs take the pixel-wise labeled images as the input data and output the generated images that meet constraints from its corresponding input images. However, these generative adversarial models have been widely studied for natural sceneries, which are not suitable for remote sensing imageries due to the complex scenes and low resolutions. Hence, the GANs perform poorly when adopted to generate remote sensing images. An aircraft recognition framework of remote sensing images based on sample generation is proposed, which consists of an improved CGAN and a recognition model, to alleviate the lack of real samples and deal with the problems mentioned above. **Method** In this framework, the masks of real aircraft images are labeled pixel by pixel. The masks of images serve as the conditions of the CGAN that are trained by the pairs of real aircraft images and corresponding masks. In this manner, the location, scale, and type of aircraft in the synthetic images can be controlled. Perceptual loss is introduced to promote the ability of the CGANs to model the scenes of remote sensing. The L2 distance between the features of real images and synthetic images extracted by the VGG-16 (visual geometry group 16-layer net) network measures the perceptual loss between the real images and synthetic images. Masked structural similarity (SSIM) loss is proposed, which forces the CGAN to focus on the masked region and improve the quality of the aircraft region in the synthetic images. SSIM is a measurement of image quantity according to the structure and texture. Masked SSIM loss is the sum of the product between masks and SSIM pixel by pixel. Afterward, the loss function of the CGAN consists of perceptual loss, masked SSIM loss, and origin CGAN loss. The recognition model in this framework is ResNet-50, which outputs the type and recognition score of an aircraft. In this paper, the recognition model trained on synthetic images is compared with the model trained on real images. The remote sensing images from QuickBird are cropped to build the real dataset, in which 800 images for each type are used for training and 1 000 images are used for testing. After data augmentation, the training dataset consists of 40 000 images, and the synthetic dataset consists of synthetic images generated by the generation module with flipped, rotated, and scaled masks. The generators are selected from different training stages to generate 2 000 synthetic images per type and determine the best end time in the training procedure. The chosen generator is used to produce different numbers of images for 10 aircraft types and find an acceptable number of synthetic images. These synthetic images serve as the training set for the recognition model, whose performances are compared. All our experiments are carried out on a single NVIDIA K80 GPU device with the framework of Pytorch, and the Adam optimizer is implemented to train the CGAN and ResNet-50 for 100 epochs. **Result** The quantities of the synthetic images from the generator with and without our proposed loss functions on the training dataset are compared. The quantitative evaluation metrics contain peak signal to noise ratio (PSNR) and SSIM. Results show that PSNR and SSIM increase by 0.88 and 0.346 using our method, respectively. In addition, recognition accuracy increases with the training epoch of the generator and the number of synthetic images. Finally, the accuracy of the recognition model trained on the synthetic dataset is 0.33% less than that of the real dataset. **Conclusion** An aircraft recognition framework of remote sensing images based on sample generation is proposed. The experiment results show that our method effectively improves the ability of CGAN to model the remote sensing scenes and alleviates the absence of data.

Key words: deep learning; convolutional neural network; generative adversarial network (GAN); optical remote sensing images; object recognition

0 引言

飞机识别是遥感影像解译领域的一项重要任务,受到了越来越多的关注。但由于成像条件、周围环境和飞机姿态的多样性,飞机识别任务具有很强的挑战性。

飞机识别算法按特征提取的方式分为传统机器学习算法和深度学习算法两类。Roopa 等人(2017)从飞机图像中提取特定的特征向量,组合后用全连接网络分类,Zhao 等人(2017)利用地标检测与模板匹配的方法实现飞机目标识别,这类传统方法存在鲁棒性差、模型容量小等问题。Wang 等人(2017)、Zhang 等人(2018)、Zuo 等人(2018)、Fu 等人(2019)利用深度卷积神经网络进行飞机识别,自动提取影像特征,此类深度学习算法识别准确率高,但其泛化性能依赖于大规模数据集。

考虑到成像条件、周围环境和飞机姿态的多样性,构建大规模飞机识别数据集难度很大。此外,实际数据中,由于各型飞机数量差异大,实际标注的数据集往往存在样本不均衡问题。数据增广是解决样本不足问题的常用方法之一,而 Song 等人(2019)的研究表明,简单的数据增广方法无法对遥感图像中的复杂场景实现高质量建模,所以该方法作用有限。

Goodfellow 等人(2014)提出的生成对抗网络(generative adversarial network, GAN)可用于数据生成任务,该模型通过学习样本的数据分布,能产生大量与真实样本分布相似的生成样本,但无法控制生成样本中目标的位置、角度和尺寸等特征。Mirza 和 Osindero(2014)对网络的输入引入“条件”,使得模型可以生成指定“条件”的仿真样本,但这种控制条件没有达到像素级的精度,因此也无法用于生成飞机样本;Isola 等人(2017)将逐像素语义标注作为生成器的输入“条件”,使得模型对生成样本的模式控制可达像素级,实现任意图到任意图的转换任务,但生成图像质量不高;Wang 等人(2018)在此基础上,提出多级生成网络和多级判别网络,以生成高分辨率的图像。这些生成对抗网络虽然可以对简单的自然场景生成逼真的图像,但对场景复杂、尺度多样且分辨率低的遥感图像的建模能力有限,导致生成样本质量低。

针对上述问题,本文提出了结合条件生成对抗

网络(conditional generative adversarial network, CGAN)样本生成的遥感图像飞机识别框架,利用基于条件生成对抗网络的样本生成方法获得数量、类别与图像模式可控的飞机识别样本集,显著减轻了样本标注的负担,且避免了样本不均衡问题。同时,针对条件生成对抗网络对遥感场景建模能力不足的问题,提出利用感知损失引入监督信息以提高生成器的建模能力,并提出基于掩膜的结构性相似性(structural similarity, SSIM)度量损失函数(masked-SSIM loss),使生成样本中的飞机目标与真实样本的结构一致性得到增强,提高了生成样本中飞机区域的图像质量。

1 飞机识别框架

1.1 框架概述

本文提出一种结合 CGAN 样本生成的遥感图像飞机识别框架,如图 1 所示, L_{PEGC} 为感知损失函数, L_{MSSIM} 为 masked-SSIM 损失函数, G 表示网络的生成器, D 表示网络的判别器。该框架包含样本生成和目标识别两部分。样本生成部分利用少量真实的飞机样本及对应的像素级标注数据对条件生成对抗网络进行训练,然后利用该网络的生成器 G 产生大量合成样本,如图 2 所示。目标识别部分利用深度残差网络 ResNet-50 作为飞机识别模型进行端到端的训练。最后用训练过的模型对真实遥感图像进行飞机类型识别。

1.1.1 条件生成对抗网络

Wang 等人(2018)提出的条件生成对抗网络由生成器 G 与判别器 D 构成。生成器用于产生与真实样本分布相似且对应条件 s 的生成样本 $G(s)$,判别器用于区分生成样本 $G(s)$ 和真实样本 x 。其损失函数为

$$\min_G \max_D L_{\text{CGAN}}(D, G) = E_{(s, x) \sim p_{\text{data}}(s, x)} [\log D(s, x)] + E_s \sim p_z(s) [\log(1 - D(s, G(s)))] \quad (1)$$

式中, $(s, x) \sim p_{\text{data}}(s, x)$ 表示样本 x 在条件 s 下服从 p_{data} 分布, $s \sim p_z(s)$ 表示输入条件 s 服从 p_{data} 分布, $E(\cdot)$ 表示期望。

该网络的生成器(图 2)由两个子网络 G_1 和 G_2 组成。 G_1 学习图像的全局语义特征,输出低分辨率样本;

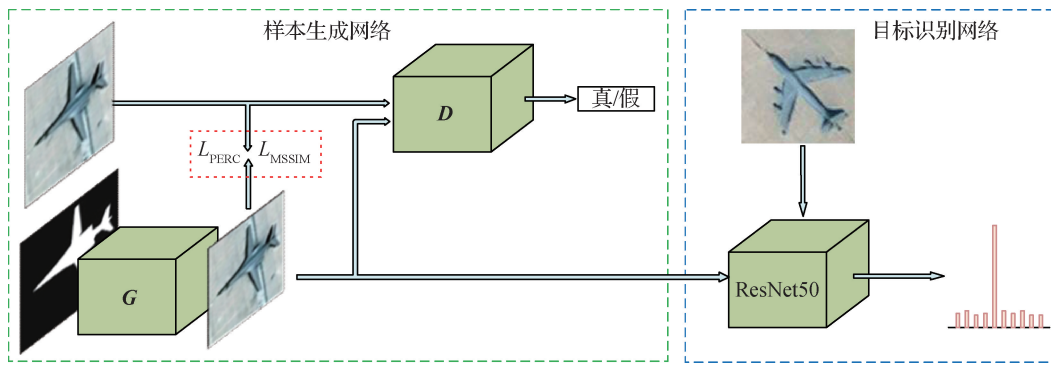


图1 结合 CGAN 样本生成的遥感图像飞机识别框架

Fig. 1 Aircraft recognition framework based on the sample generation from CGAN

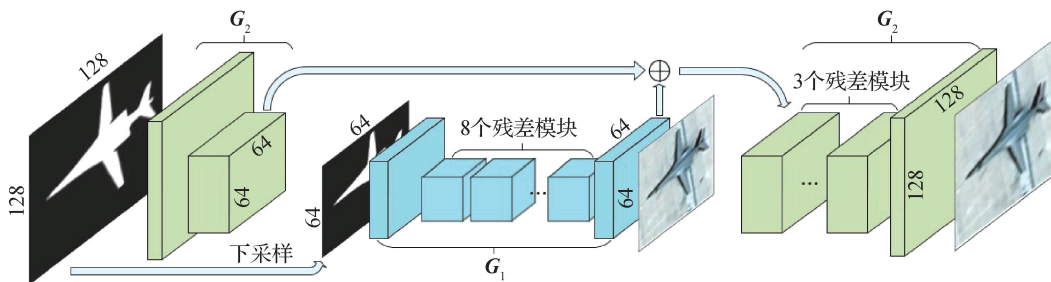


图2 样本生成模块中生成器的网络结构

Fig. 2 Generator network in the sample generation module

G_2 学习图像的局部特征,输出比 G_1 分辨率高一倍的生成样本,二者相互协作,输出高质量的生成样本。

该模型对自然场景可以产生高质量的生成样本,但是对场景复杂、分辨率低、尺度多样、前景与背景高度耦合的遥感图像建模能力不足,使得生成的遥感图像质量受限。

1.1.2 识别网络

为了解决深层神经网络训练过程中的梯度消失和梯度爆炸问题,He 等人(2016)提出了残差神经网络,使得深层神经网络的训练变得可行,且取得了较好的性能,因此得到广泛应用。该网络的核心模块是残差块(residual block),结构如图3所示,主要由跳跃连接 x 和残差连接 $F(x)$ 构成,其原理为

$$y = F(x) + x \quad (2)$$

ResNet-50 网络由 4 组残差块通过级联构成,共有 50 个卷积层,具有很强的特征提取能力且运行速度较快。用该网络构建飞机类型识别模型(图1右侧虚线框),首先使用 ResNet-50 网络提取图像的特征,然后利用全连接层进行分类,最后通过 softmax 激活函数输出图像中的飞机属于各类型的概率值,概率值最大的类型即预测出的飞机类型。

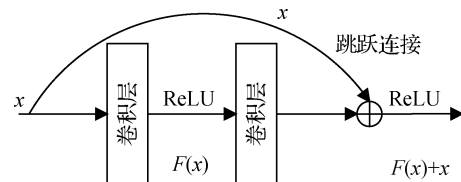


图3 残差块的结构

Fig. 3 Structure of the residual block

1.2 感知损失

遥感图像中的背景与前景在图像模式上高度耦合,所以生成模型必须具备对这种复杂场景的建模能力。Wang 等人(2018)提出的条件生成对抗网络的生成器只有图像块级别的判别损失,而没有高维特征级别的约束,导致生成样本与真实样本在空间和语义上相似性差,因而无法很好地学习遥感图像中的复杂场景。

为了解决这个问题,利用感知损失提高生成器对遥感图像的建模能力。感知损失由 Johnson 等人(2016)提出,并在提高图像分辨率任务上取得了显著效果。该方法用 ImageNet(Russakovsky 等,2015)的分类任务对 VGG-16 网络(visual geometry group

16-layer net) (Simonyan 和 Zisserman, 2015) 进行预训练, 然后用预训练过的网络作为损失网络 ϕ 提取图像的特征 (如图 4 所示), 并计算生成样本与目标样本在特征图上的距离, 通过缩小这种距离, 使两者在深度卷积神经网络上表现的语义分布更为相似。其目标函数 L_{PERC} 表示为

$$L_{\text{PERC}}(\mathbf{x}, \mathbf{G}(s)) = E_{(x,s)} \left[\frac{1}{C_j H_j W_j} \|\phi_j(\mathbf{x}) - \phi_j(\mathbf{G}(s))\|_2^2 \right] \quad (3)$$

式中, ϕ_j 表示 VGG-16 网络第 j 层 (一般选择第 3 个 block 的第 3 层) 特征, 最后将宽 (W_j)、高 (H_j) 和通道 (C_j) 3 个维度取平均值作为生成样本 $\mathbf{G}(s)$ 与真实样本 \mathbf{x} 的感知损失 (perceptual loss)。

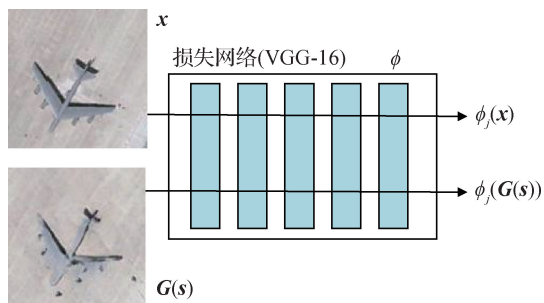


图 4 感知损失示意图

Fig. 4 Perceptual loss

该方法利用损失网络缩小生成样本与目标样本在语义分布上的距离, 而不是简单地使模型直接学习目标样本的像素值, 生成的样本图像细节更丰富、图像纹理更自然。因此, 用感知损失函数对生成器添加特征级别的监督信息, 提高生成样本与真实样本的语义相似性, 从而使得生成的飞机样本图像更接近真实图像。

1.3 基于目标掩膜的结构相似性度量损失函数

在飞机识别任务中, 结构和纹理是区分不同飞机类型的关键因素之一, 为了保证飞机识别模型从生成样本中学习到的飞机结构和纹理特征与真实样本一致, 必须使生成图像在飞机结构和纹理上与真实图像尽可能接近。原始的生成对抗网络无法针对飞机目标区域实现高质量的建模, 而感知损失函数可以使生成图像在特征和语义上与真实样本相近, 但无法保证生成图像中飞机目标区域在结构和纹理等细节上与真实图像一致。

为了提高生成样本中飞机区域与真实样本的一致性, Wang 等人 (2004) 提出利用结构相似性

(SSIM) 度量构建损失函数, 对生成器加入监督信息。结构相似性度量是一种用于评估图像间相似度的指标, 利用滑窗的方式计算图像间局部的相似度, 最后计算全局的结构相似度, 其计算过程为

$$l(\mathbf{X}, \mathbf{Y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (4)$$

$$c(\mathbf{X}, \mathbf{Y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (5)$$

$$d(\mathbf{X}, \mathbf{Y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (6)$$

$$S(\mathbf{X}, \mathbf{Y}) = [l(\mathbf{X}, \mathbf{Y})]^\alpha \cdot [c(\mathbf{X}, \mathbf{Y})]^\beta \cdot [d(\mathbf{X}, \mathbf{Y})]^\gamma \quad (7)$$

$$C_1 = (K_1L)^2, C_2 = (K_2L)^2, C_3 = C_2/2 \quad (8)$$

式中, \mathbf{X} 和 \mathbf{Y} 表示两幅对齐的图像或图像块, μ_x 、 σ_x^2 和 σ_{xy} 分别表示图像 \mathbf{X} 的均值、方差以及图像 \mathbf{X} 与图像 \mathbf{Y} 的协方差, $l(\mathbf{X}, \mathbf{Y})$ 、 $c(\mathbf{X}, \mathbf{Y})$ 和 $d(\mathbf{X}, \mathbf{Y})$ 分别表示图像 \mathbf{X} 与 \mathbf{Y} 的亮度、对比度以及结构的距离度量, $S(\mathbf{X}, \mathbf{Y})$ 为两幅图像的 SSIM (其中 $\alpha = \beta = \gamma = 1$), L 表示图像的像素值范围 ($0 \sim 255$), 系数 $K_1 = 0.01$ 且 $K_2 = 0.03$, 因此 SSIM 计算过程可简化为

$$S(\mathbf{X}, \mathbf{Y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (9)$$

SSIM 取值为 $[0, 1]$, 值越大表示图像越相似。直接使用该指标作为损失函数, 生成器会同时学习前景和背景, 导致生成图像的背景区域也与真实图像一致, 破坏了背景区域的多样性。因此, 提出基于目标掩膜的结构相似性度量损失函数。在计算生成图像与真实图像的 SSIM 损失时, 利用飞机目标的掩膜 M_{ij} 屏蔽背景区域, 只保留飞机目标所在区域的 SSIM 损失。利用该损失函数, 可以提高生成图像 $\mathbf{G}(s)$ 中的飞机区域与真实图像 \mathbf{x} 在结构和纹理上的一致性, 其原理为

$$L_{\text{MSSIM}}(\mathbf{x}, \mathbf{G}(s)) = E_{(x, G(s))} \times \left[\frac{1}{\sum_j^H \sum_i^W M_{ij}} \sum_j^H \sum_i^W M_{ij} \cdot (1 - S_{ij}(\mathbf{x}_{ij}, \mathbf{G}(s)_{ij})) \right] \quad (10)$$

$$M_{ij} = \begin{cases} 1 & m_{ij} > 0 \\ 0 & m_{ij} = 0 \end{cases} \quad (11)$$

式中, H 和 W 表示图像的高和宽, S_{ij} 指位于图像像素 (i, j) 处的局部结构相似性, M_{ij} 表示图像的掩

膜,飞机所在的区域取值为1,其他背景区域取值为0,最后取该飞机区域 S_j 的平均值。该方法实现了只计算飞机区域的结构一致性损失,而对背景区域不产生影响。

1.4 CGAN 的损失函数

条件生成对抗网络的损失函数由原始的损失函数 $L_{CGAN}(D, G)$ 、感知损失函数 $L_{PERC}(x, G(s))$ 和基于目标掩膜的结构相似性度量损失函数 $L_{MSSIM}(x, G(s))$ 构成,即

$$L(x, s, D, G) = L_{CGAN}(D, G) + \lambda L_{PERC}(x, G(s)) + \mu L_{MSSIM}(x, G(s)) \quad (12)$$

式中, λ 和 μ 分别是感知损失和 masked-SSIM 损失的权重,分别取 0.5 和 1。

2 数据处理

2.1 飞机目标掩膜的标注方法

在遥感图像中,飞机目标具有刚体特性,其形状和内部各部件的相对位置不变,飞机的掩膜图像可以有效表征飞机的类型信息。因此,使用飞机的掩膜图像作为条件生成对抗网络输入的“条件”,以控制生成的飞机样本的模式,生成指定内容的样本。

图 5 是 10 类飞机的掩膜图像。标注图像时,在目标区域(图 5 白色部分)内填充飞机类型号(取值为 [1, 10] 的整数),在背景区域(图 5 黑色部分)填充 0。与 Wang 等人(2018)的方法一样,为了便于模型训练,将掩膜图像沿着通道维度作 one-hot 编码,然后将其输入网络。

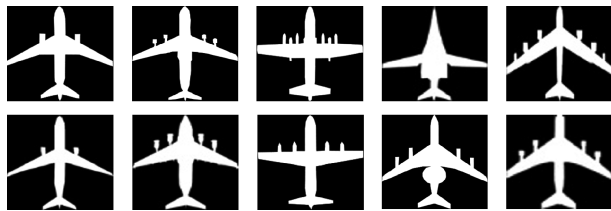


图 5 10 类飞机的掩膜图像

Fig. 5 The masks of aircrafts of ten types

2.2 真实样本数据集的获取方法

从大幅遥感图像上,裁切出大量只包含单个飞机目标的切片,以构建真实样本数据集。为保证模型能适应不同尺度和不同分辨率下的目标,在一定范围内调整裁切窗口,并将每个切片缩放(二次线性插值)至 128×128 像素。在获取切片后,为每个

切片标注对应的目标掩膜。为保证数据的多样性,对每个裁切窗口的中心位置沿着 x 和 y 方向添加随机扰动,使飞机目标尽可能出现在图像切片的不同位置。在此基础上,对获取的切片作任意角度随机旋转,切片对应的掩膜图像也作对应处理,从而实现样本扩充,构建出真实样本的数据集。

2.3 生成样本数据集的获取方法

条件生成对抗网络训练完成后,使用该网络的生成器产生一定数量的生成样本。首先,对每一类飞机目标的掩膜进行随机缩放(最近邻插值)、旋转和平移等预处理;然后,将处理后的掩膜作为生成器的输入,并得到对应的生成样本。对每种飞机类型都按照这种方法产生一定数目的样本,从而构建生成样本的数据集。

3 实验与分析

3.1 数据集

为构建真实样本数据集,从 QuickBird 卫星影像数据中获取各国机场的大幅全色光学遥感图像,图像分辨率为 0.61 m。通过裁切的方式获取大量只包含单个飞机的切片,裁切窗口的变化范围为 [96, 192],图像中心以目标中心为基准,扰动范围为 [40, 40],裁切后的切片缩放至 128×128 像素。

通过上述操作,为每类飞机获取了 800 幅切片图像,10 类共计 8 000 幅。再通过样本扩充,共获得 40 000 幅切片图像,并以此构建真实样本的训练集,用于训练条件生成对抗网络和飞机目标识别网络。此外,选取与前述数据在地理空间上不相交的区域,按照相同操作,为每类飞机获取了 1 000 幅切片图像,10 类共 10 000 幅,并以此构建真实样本的测试集,用于评估不同的识别模型的性能。

3.2 模型参数设置

实验使用 Adam 算法(Kingma 和 Ba, 2017)对条件生成对抗网络进行训练, batch 大小为 16, 训练 100 个周期。目标识别网络同样使用 Adam 算法进行训练,训练 100 个周期。上述算法均使用 Pytorch(Paszke 等, 2019)实现,并在一块 NVIDIA Tesla K80 显卡上进行训练。使用识别准确率(accuracy)评估模型对飞机目标的识别性能,该值越大表示识别性能越好,其计算过程为

$$R(TP, P, N, TN) = \frac{TP + TN}{P + N} \quad (13)$$

式中, TP 为各类别被正确识别的正样本数量, TN 为各类别被正确识别的负样本数量, P 为各类别的正样本数量, N 为各类别的负样本数量。

用峰值信噪比 (peak signal to noise ratio, PSNR) 和结构相似性度量 (SSIM) 指标来量化生成样本的图像质量, 以评估所提出的损失函数对生成器性能的影响。

3.3 损失函数

实验对比了加入提出的损失函数前后的生成器产生的图像与对应的真实图像的相似性, 定量的评估结果如表 1 所示。可以看出, 加入感知损失后, 生成图像与真实图像的 PSNR 得到了明显提升。由感知损失引入的特征级别的监督信息使得生成模型产生的样本与真实样本的数据分布更接近, 对应的像素分布的差异也进一步缩小, 而 PSNR 主要统计对应像素的距离, 所以 PSNR 得到提高。此外, 还加入了提出的 masked-SSIM 损失函数, 使得模型朝着 SSIM 更大的方向进行优化, 所以加入该损失函数后 SSIM 得到了显著的提高。

表 1 加入损失函数前后生成样本的图像质量

Table 1 The quality of the generated images against the generator with/without loss function

方法	PSNR/dB	SSIM
CGAN	17.23	0.352
CGAN + L_{PERC}	18.02	0.446
CGAN + L_{PERC} + L_{MSSIM}	18.11	0.698

注: 加粗字体为每列最优结果。

图 6 为实验获取的部分样本图像。其中, 第 2 行的生成图像由原始条件生成对抗网络产生, 图像中飞机的引擎、机尾等部件出现缺失, 且背景区域纹理单一, 与真实样本差异大。加入感知损失后, 生成样本如图 6 第 3 行所示, 其背景区域的纹理更加复杂, 较好还原了真实样本的整体分布情况, 表明使用的感知损失函数有利于提高模型对图像背景区域的建模能力。但这些样本中飞机引擎边缘及机身纹理与真实图像仍有差距。加入 masked-SSIM 损失函数后, 在生成图像 (图 6 第 4 行) 中的这种问题得到明显改善, 表明该损失函数提高了生成模型对飞机的结构和纹理的建模能力。

定量与定性的实验结果表明, 这两种损失函数分别提高了模型对背景区域的纹理、前景目标的结构和纹理的建模能力, 从而生成了质量更高的样本。

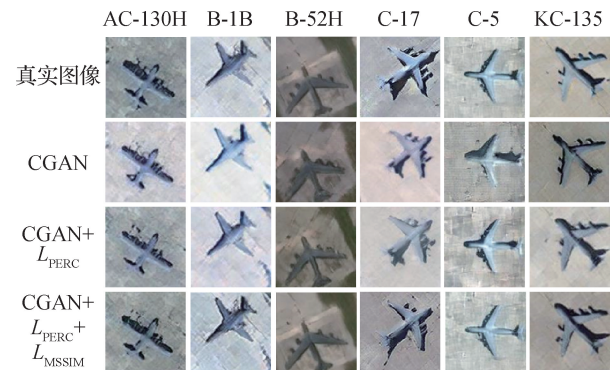


图 6 CGAN 加入损失函数前后的生成样本

Fig. 6 The images generated by the CGAN with and without proposed loss

3.4 生成样本的质量

为了更有效地获取生成样本, 实验定量地研究了生成样本对识别准确率的影响。首先分别选取了训练周期数为 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 的条件生成对抗网络, 然后用这些模型分别为这 10 种飞机逐类随机生成 2 000 幅样本图像, 即每个模型均产生 20 000 幅生成样本图像, 如图 7 所示。用这些生成样本训练识别模型, 并使用真实样本的测试集评估其识别准确率, 测试结果如图 8 所示。

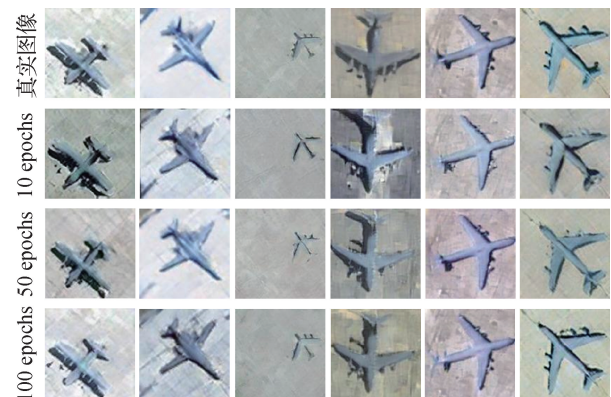


图 7 不同训练周期下的生成样本

Fig. 7 Samples generated by CGAN after different training epochs

从图 7 可以看出, 随着 CGAN 训练次数增加, 生成图像越来越逼真, 语义越来越丰富, 即生成样本的质量在逐步提高。同时, 识别模型的准确率也随之提高 (图 8)。显然, 提高生成样本的质量能有效提

高识别准确率。

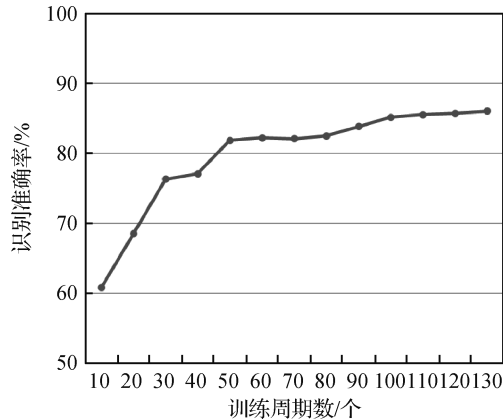


图8 识别准确率与CGAN的训练周期数

Fig. 8 Recognition accuracy against training epochs of CGAN

在CGAN训练10个周期后,识别模型的准确率仅为59.64%。随着训练次数增加,准确率迅速提高。在生成器训练50个周期后,准确率的提高速度逐渐减缓,主要原因是生成器的损失函数趋于稳定,此时生成样本的质量不再显著提高。在80个周期后,学习率随之衰减到更小的值,有助于CGAN的进一步训练,准确率再次提高。到了100个周期后,损失函数保持稳定,学习率继续衰减,训练过程逐渐稳定,识别准确率不再具有明显的增长趋势。考虑到训练成本与带来的性能增益之间的权衡,均选取第100个周期的实验结果。

3.5 生成样本的数量

生成对抗网络的一种优势是可以生成大量的样本以打破数据规模的限制,但选取生成样本的数量却需要进一步研究。因此,实验对比了不同数量的生成样本对识别模型性能的影响。

根据3.4节的实验结果,使用训练100个周期的生成器获取生成样本,并取得了较好的识别性能。因此,实验使用该模型分别为每类飞机产生2 000, 4 000, 8 000, 12 000, 20 000, 40 000幅样本图像用于训练识别模型,并在真实样本测试集上评估其性能,结果如表2所示。

实验结果表明,随着样本数量的增加,样本的多样性逐渐增加,识别模型的过拟合程度有效减轻,因此识别准确率有了大幅提升。在每类40 000幅样本图像的情况下,识别准确率比每类2 000幅样本图像的情况提高了7.07%。本实验结果表明,增加生成样本的数量能有效提高识别准确率。

表2 不同样本数量下的目标识别准确率

Table 2 The accuracy against the number of images

每类样本数量/个	识别准确率/%
2 000	85.14
4 000	86.39
8 000	88.42
12 000	89.17
20 000	91.33
40 000	92.21

3.6 生成样本与真实样本的目标识别性能

本文将提出的识别框架用生成样本替代真实样本训练飞机识别模型,为了评估这种方法的有效性,用训练了100个周期的CGAN产生40万个样本以构建生成样本的训练集,同时取真实样本构建真实训练集,然后分别用这两种数据集训练识别模型,并在真实样本的测试集上评估其识别准确率,定量的测试结果如表3所示,所提出的飞机识别框架的识别混淆矩阵如图9所示。

表3 不同数据源的目标识别准确率

Table 3 The recognition accuracy of samples from different data sources

数据源	识别准确率
真实样本	92.54
生成样本	92.21

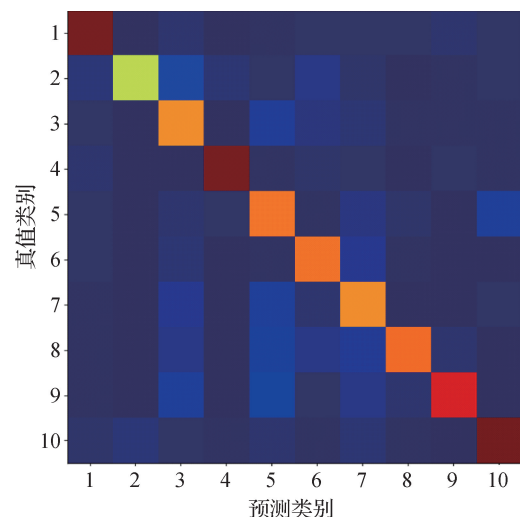


图9 由生成样本训练的识别模型的混淆矩阵

Fig. 9 The confusion matrix of recognition model trained by the generated samples

实验结果表明,与用真实样本训练的识别模型的准确率相比,用生成样本训练的识别模型的准确率低 0.33%,两者性能相近。因此使用生成样本可以有效代替真实样本训练飞机目标识别模型。

实验通过采用少量的真实样本模拟真实飞机样本不足的情况。用于训练识别模型的真实样本为每类 800 幅,表 3 中的生成样本为每类 40 000 幅;而用于评估识别准确率的真实测试样本为每类 1 000 幅。表 3 表明,在仅有少量真实样本的情况下,通过构建大规模高质量的生成样本数据集,可以达到与真实样本相近的准确率,因此可以有效解决飞机识别任务中样本不足的问题。

3.7 讨论分析

生成样本与真实样本训练的识别模型存在性能差异。真实的训练样本与测试样本一般具有相同的分布特性,而生成样本的分布与真实样本的分布存在一定差异,原因是真实样本中所含的独特语义信息无法被条件生成对抗网络完全学习和表达。因而在训练数据规模相同的情况下,用生成样本训练的识别模型性能较差。但实验结果也同时表明,可以通过提高生成样本的质量和增加训练数据规模的方式进行弥补,使其达到相近的识别性能。

图 10 显示质量差的生成样本与识别错误的真实样本。虽然所提方法使得生成模型的性能得到了有效提升,但是仍然有少量生成样本的质量较差,如图 10 第 1 行所示。一方面,这些质量差的图像分辨率较低,可能是由真实样本中存在的低分辨率图像造成,如图 10 第 3 行所示。另一方面,这些生成样本质量差主要体现在其图像结构内部存在一定的棋盘效应,这种问题主要来自生成器模块(图 2)中的反卷积操作。反卷积很容易出现“不均匀的重叠”,当内核大小(输出窗口大小)不能被步幅(顶部各点之间的间距)整除时,反卷积具有不均匀的重叠,从而导致在分辨率放大的过程中,产生不清晰的棋盘效应问题。同时,在计算卷积层梯度的过程中,反卷积的梯度反向传播过程也会导致一定的渐变棋盘效应(Odena 等,2016)。

虽然识别模型得到了很高的准确率,但是仍有一些飞机目标识别错误,部分样本如图 10 第 2 行所示。选用的识别模型为 ResNet-50,受该模型识别准

确率的限制,部分样本识别错误。此外,真实样本中存在分辨率较低或尺寸较小的飞机样本(图 10 第 3 行),导致生成样本也出现此种情况,而这些样本缺少更具判别性的特征,使得模型产生误判。因此,在未来的工作中,将从提高遥感图像的分辨率方面进行进一步的研究。

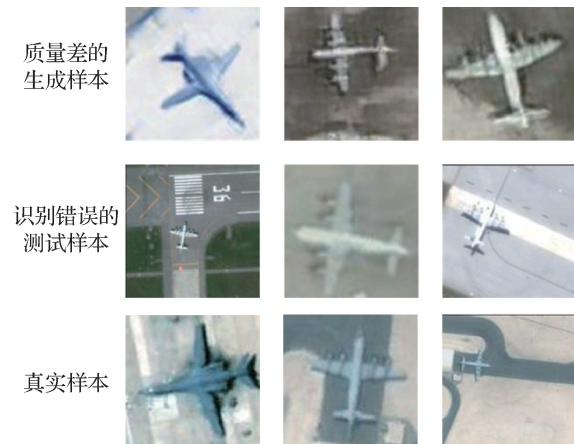


图 10 质量差的生成样本与识别错误的真实样本

Fig. 10 The synthetic images with low quality and the misidentified real images

4 结论

针对遥感图像解译领域中基于深度神经网络的飞机识别模型对训练数据需求量大但可获得的数据有限以及条件生成对抗网络对遥感场景建模能力有限、生成样本质量低的问题,提出了一个结合 CGAN 样本生成的遥感图像飞机识别框架。该框架利用感知损失和提出的基于掩膜的 SSIM 损失增强条件生成对抗网络中生成器对复杂的遥感场景的建模能力,使用改进后的网络所生成的样本代替真实样本训练飞机识别模型。实验结果显示,该方法提高了生成样本与真实样本的结构、纹理和语义一致性,获得了质量更高的生成样本。利用这些样本构建的大规模高质量生成样本数据集可以有效替代真实样本对飞机识别模型进行端到端训练,从而有效地解决了大规模真实飞机样本数据集难以构建的难题。由于真实样本中存在一些质量差、分辨率低以及目标尺寸小的样本,导致生成样本中也出现了一些质量差的图像,使得飞机识别模型出现识别错误。在未来的研究中,将进一步研究提高遥感图像分

分辨率的方法,以获取更高质量的真实图像,从而改善生成样本的质量并提高识别性能。该方法仅在飞机样本生成方面进行了改进,生成样本质量还有很大的提升空间。此外,更适合飞机识别的网络模型也有待研究。

致谢 实验的计算平台、数据获取及标注工作得到了中国科学院空天信息创新研究院网络信息体系技术重点实验室的大力支持,在此表示感谢。

参考文献 (References)

- Fu K, Dai W, Zhang Y, Wang Z R, Yan M L and Sun X. 2019. Multi-CAM: multiple class activation mapping for aircraft recognition in remote sensing images. *Remote Sensing*, 11(5): #544 [DOI: 10.3390/rs11050544]
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2014. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press: 2672-2680
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Isola P, Zhu J Y, Zhou T J and Efros A A. 2017. Image-to-Image translation with conditional adversarial networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5967-5976 [DOI: 10.1109/CVPR.2017.632]
- Johnson J, Alahi A and Li F F. 2016. Perceptual losses for real-time style transfer and super-resolution//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer: 694-711 [DOI: 10.1007/978-3-319-46475-6_43]
- Kingma P and Ba J. 2017. Adam: a method for stochastic optimization [EB/OL]. [2019-12-08]. <https://arxiv.org/pdf/1412.6980.pdf>
- Mirza M and Osindero S. 2014. Conditional generative adversarial nets [EB/OL]. [2019-12-08]. <https://arxiv.org/pdf/1411.1784.pdf>
- Odena A, Dumoulin V O and Olah C. 2016. Deconvolution and checkerboard artifacts [EB/OL]. [2019-12-08]. <http://distill.pub/2016/deconv-checkerboard>
- Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z M, Gimelshein N, Antiga L, Desmaison A, Köpf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J J and Chintala S. 2019. PyTorch: an imperative style, high-performance deep learning library//Proceedings of the 33rd Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates: 8026-8037
- Roopa K, Rama Murthy T V and Cyril Prasanna Raj P. 2017. Neural network classifier for fighter aircraft model recognition. *Journal of Intelligent Systems*, 27(3): 447-463 [DOI: 10.1515/jisys-2016-0087]
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S A, Huang Z H, Karpathy A, Khosla A, Bernstein M, Berg A C and Li F F. 2015. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3): 211-252 [DOI: 10.1007/s11263-015-0816-y]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-12-08]. <https://arxiv.org/pdf/1409.1556.pdf>
- Song J, Gao S H, Zhu Y Q and Ma C Y. 2019. A survey of remote sensing image classification based on CNNs. *Big Earth Data*, 3(3): 232-254 [DOI: 10.1080/20964471.2019.1657720]
- Wang H Z, Gong Y C, Wang Y, Wang L F and Pan C H. 2017. DeepPlane: a unified deep model for aircraft detection and recognition in remote sensing images. *Journal of Applied Remote Sensing*, 11(4): #042606 [DOI: 10.1117/1.JRS.11.042606]
- Wang T C, Liu M Y, Zhu J Y, Tao A, Kautz J and Catanzaro B. 2018. High-resolution image synthesis and semantic manipulation with conditional GANs//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8798-8807 [DOI: 10.1109/CVPR.2018.00917]
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]
- Zhang Y H, Sun H, Zuo J W, Wang H Q, Xu G L and Sun X. 2018. Aircraft type recognition in remote sensing images based on feature learning with conditional generative adversarial networks. *Remote Sensing*, 10(7): #1123 [DOI: 10.3390/rs10071123]
- Zhao A, Fu K, Wang S Y, Zuo J W, Zhang Y H, Hu Y F and Wang H Q. 2017. Aircraft recognition based on landmark detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 14(8): 1413-1417 [DOI: 10.1109/LGRS.2017.2715858]
- Zuo J W, Xu G L, Fu K, Sun X and Sun H. 2018. Aircraft type recognition based on segmentation with deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(2): 282-286 [DOI: 10.1109/LGRS.2017.2786232]

作者简介



王耀领, 1984年生, 男, 博士研究生, 主要研究方向为遥感图像理解。

E-mail: 13311262527@163.com



王宏琦, 通信作者, 男, 研究员, 博士生导师, 主要研究方向为遥感图像处理、目标检测识别及相关领域。

E-mail: wiecas@sina.com

许滔, 男, 博士研究生, 主要研究方向为遥感图像处理、目标检测识别及相关领域。E-mail: xutao17@mails.ucas.ac.cn