

# 少样本条件下基于生成对抗网络的 遥感图像数据增强

姜雨辰\*, 朱斌

国防科技大学电子对抗学院脉冲功率激光技术国家重点实验室, 安徽 合肥 230009

**摘要** 针对小样本条件下遥感图像目标的检测精度受到卷积神经网络过拟合影响的问题, 提出一种基于生成对抗网络的数据增强方法, 利用判别模型为生成模型同时提供图像的局部决策与全局决策, 以提高生成模型生成图像的质量, 并将生成的目标与训练集图像进行融合得到新的样本, 且新生成的样本不需人工标注。实验结果表明, 在原始数据中加入生成数据后, 检测识别精度有所提高, 且本文方法与基于图像仿射变换的数据增强方法的叠加使用进一步提高了数据增强的效果。

**关键词** 图像处理; 数据增强; 生成对抗网络; 判别模型; 生成模型

中图分类号 TP391.9

文献标志码 A

doi: 10.3788/LOP202158.0810022

## Data Augmentation for Remote Sensing Image Based on Generative Adversarial Networks Under Condition of Few Samples

Jiang Yuchen\*, Zhu Bin

State Key Laboratory of Pulsed Power Laser Technology, College of Electronic Countermeasures,  
National University of Defense Technology, Hefei, Anhui 230009, China

**Abstract** To solve the problem that the detection accuracy of remote sensing image targets is affected by convolution neural network overfitting under the condition of small samples, a data augmentation method based on generative adversarial networks is proposed. The discrimination model is used to provide local and global decisions for the generation model to improve the quality of the image generated by the generative model. The new samples are obtained by fusing the generated target and the training set image, and the new samples do not need to be labeled manually. Experimental results show that: the accuracy of detection and recognition is improved after adding the generated data to the original data; this method can be superimposed with the data augmentation method based on image affine transformation to further improve the effect of data augmentation.

**Key words** image processing; data augmentation; generative adversarial networks; discrimination model; generative model

**OCIS codes** 100.2980; 100.3008; 150.1135

## 1 引言

随着遥感技术的迅猛发展, 遥感图像目标检测识别在各领域的应用极为广泛, 其中卷积神经网络常作为一种重要的研究方法。卷积神经网络

优异的性能、强大的特征提取能力是建立在丰富的数据集基础之上的, 数据量不足时会导致过拟合, 进而影响算法性能。然而, 在现实世界中, 受数据记录条件的限制, 大量有标签的遥感图像数据的获得通常是不现实的, 人们往往仅得到少量

收稿日期: 2020-07-30; 修回日期: 2020-09-13; 录用日期: 2020-09-22

\* E-mail: jyc2647118@126.com

带标签的数据样本。实际应用中出现了因某些领域的数据集中样本数量过少而限制现有检测识别算法性能的情况。

针对如何在少样本条件下减小过拟合的影响,当前常用的数据增强方法主要有图像仿射变换、信息删除、图像融合、生成模型<sup>[1-7]</sup>等方法。图像仿射变换方法主要指对图像进行平移、尺度变换、对比度变换、噪声扰动等方法,文献[8]中将图像仿射变换中的各项操作进行离散优化以应用于图像分类问题。信息删除主要指删除图像中的一部分内容,其中较为典型的信息删除方法包括剪切(Cutout)、网格掩码(Gridmask)<sup>[9-10]</sup>等。图像融合的方法主要指将两张或多张图片融合为一张图片的方法,其中较为典型的方法有混合(Mixup)、自动混合(Augmix)<sup>[11-12]</sup>等。利用生成模型的方法生成数据,常用的生成模型主要包括变分自编码器(VAE)、生成对抗网络(GAN)<sup>[13-14]</sup>等,这类方法比一般的图像预处理方法更为复杂,但其生成的样本也更加多样。

本文对基于样式的生成对抗网络(StyleGAN2)<sup>[15]</sup>进行改进并将其应用于少样本条件下的遥感图像数据增强中,在经过数据增强后的数据集上利用 yolov3 目标检测算法<sup>[16]</sup>进行检测,检测精度有所提高。本文主要的工作包括 3 个方面:1)提出了一种基于生成对抗网络的遥感图像数据增强流程;2)为提高图像生成质量,综合图像局部信息与全局信息,对网络判别模型进行改进,将其修改为 U 形网络(U-net)<sup>[17]</sup>,并对网络损失函数进行修改;3)利用检测模型的检测精度作为评价指标,评价了利用改进后的 StyleGAN2 网络进行数据增强的效果,并将其与图像仿射变换方法进行了比较。

## 2 生成对抗网络

GAN 的基本结构框架图如图 1 所示。生成对抗网络中设定了一个生成模型 G 和一个判别模型 D,其中生成模型负责生成接近真实数据的生成数据,而判别模型试图区分真实数据与生成模型创造出的伪数据。生成模型 G(其输出为  $G$ )以噪声信号 noise 作为输入变量,将其映射到生成模型的数据空间,从而得到生成模型的概率分布。判别模型 D(其输出为  $D$ )的输入为真实数据(real data)和来自生成模型 G 的生成数据(fake data),通过输出一个标量来反映输入来自真实数据而不是生成数据的概率。

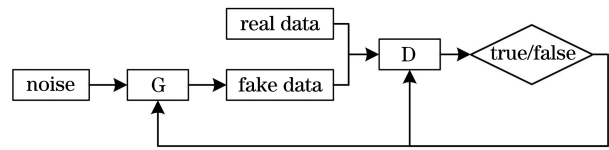


图 1 GAN 的基本结构框架图

Fig. 1 Basic structural frame of GAN

GAN 中通过训练 D 使其最大化,判断真实数据和生成数据的概率,同时训练 G 使其最小化  $\ln\{1 - D[G(z)]\}$ ,它的优化函数为

$$\min_G \max_D V(D, G) = E_x[\ln D(x)] + E_z \ln\{1 - D[G(z)]\}, \quad (1)$$

式中: $V(D, G)$ 表示目标函数; $E_x$ 、 $E_z$ 分别表示真实数据  $x$  和噪声数据  $z$  的数学期望; $G(z)$ 表示生成模型的输出; $D(x)$ 、 $D[G(z)]$ 分别表示判别模型对应真实数据与生成模型生成的伪数据的输出。

判别模型 D 的梯度为

$$\nabla_D = \nabla_{\theta_D} \frac{1}{m} \sum_{i=1}^m \{ \ln D(x^{(i)}) + \ln\{1 - D[G(z^{(i)})]\} \}, \quad (2)$$

生成模型 G 的梯度为

$$\nabla_G = \nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m \{ 1 - \ln\{1 - D[G(z^{(i)})]\} \}, \quad (3)$$

式中: $\nabla_{\theta_D}$ 、 $\nabla_{\theta_G}$ 分别表示对判别模型、生成模型求梯度; $m$ 表示批样本数量; $i$ 、 $j$ 表示样本序号; $x$ 、 $z$ 分别表示真实数据与噪声。

最终二者达到纳什均衡,判别器输出为 0.5,即判别模型 D 无法判断输入是来自真实数据还是来自生成数据。

随着近些年来研究人员对生成对抗网络的不断改进,其生成图像的质量也不断提高,如文献[18]中提出了深度卷积生成对抗网络(DCGAN),该网络使用了跨距卷积的方法,并将卷积神经网络与生成对抗网络结合起来。文献[19]中提出了基于 Wasserstein 生成对抗网络(WGAN)的模型,该模型使用 Wasserstein 距离来代替 Jensen-Shannon(JS)散度评估实际样本与生成样本之间的距离,相较于原始 GAN 进一步提高了训练的稳定性。文献[20]中提出了基于样式的生成对抗网络 StyleGAN,使用潜变量并经过空间映射得到中间变量,将该过程作为 style 控制生成模型的生成过程,StyleGAN2 针对 StyleGAN 生成图片中出现的伪影进行了改进,主要是将实例正则化的过程改进为调制-解调的过程。

### 3 遥感图像数据增强流程与算法

#### 3.1 遥感图像数据增强流程

与图像仿射变换、信息删除、图像融合等数据增强方法不同,本文通过训练生成对抗网络生成不同于原始图像的异构图像,并将生成图像与背景进行融合,得到新样本。具体流程为:1)从原始图像中截取目标,用于训练生成对抗网络;2)利用生成对抗网络生成目标样本;3)将生成目标与背景图片进行融合,得到融合后的生成样本。遥感图像数据增强流程图如图 2 所示。

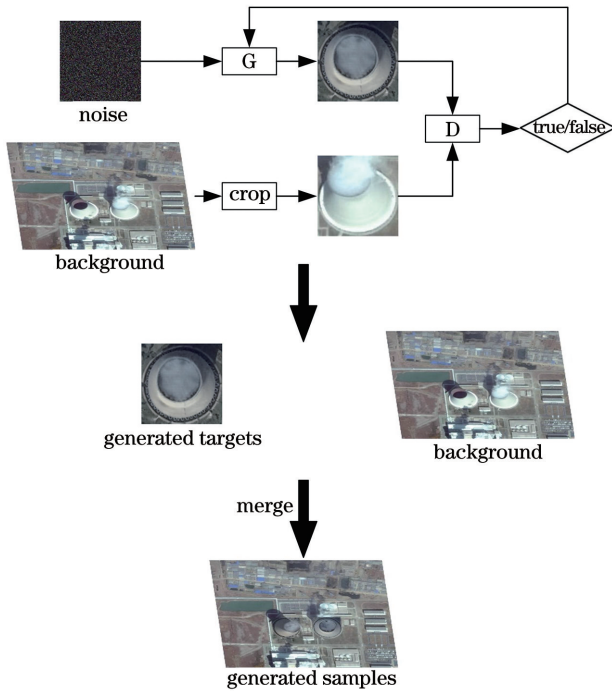


图 2 遥感图像数据增强流程图

Fig. 2 Data augmentation process of remote sensing image

#### 3.2 基于 StyleGAN2 模型的改进

这里在 StyleGAN2 的基础上对生成对抗网络进行改进,重点对判别模型的结构进行了改进,将其调整为 U-Net 结构,使它同时输出属于真实类或伪类图像的全局和局部决策。其中以原 StyleGAN2 的判别模型作为编码器模块执行全局决策,解码器模块类似生成模型,输出每像素类决策,并向生成模型提供空间相干反馈。这种体系结构的改变使得判别模型的能力得到增强,它鼓励生成模型生成细节更为丰富的图像,也使得生成模型更难以欺骗判别模型,从而提高生成样本的质量。判别模型结构如图 3 所示,图中列出了判别模型中的三层结构,且省略了卷积模块中的跳跃连接。

图 3 中 Conv 表示卷积核为  $3 \times 3$  的卷积,

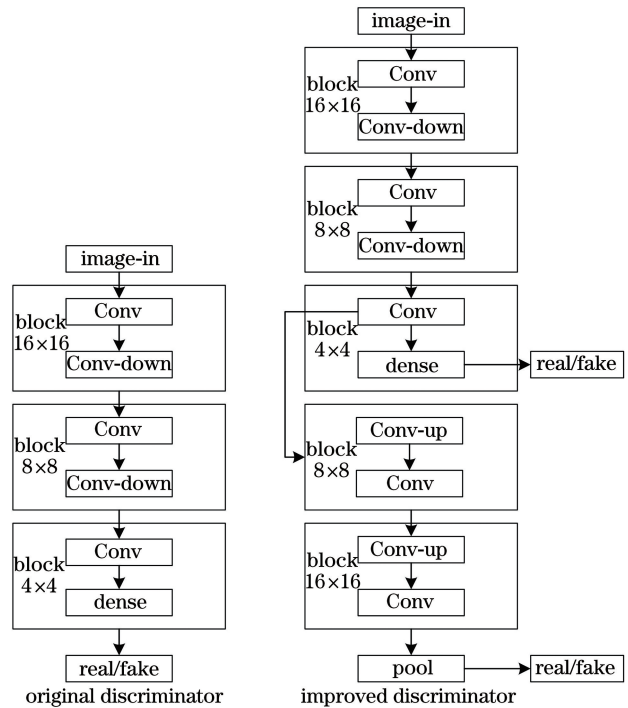


图 3 判别模型结构图

Fig. 3 Structural diagram of discrimination model

Conv-down 表示步长为 2 的卷积(用于降采样), Conv-up 表示上采样后卷积, dense 表示全连接层, pool 表示平均池化层。U-net 的编码器部分从 dense 层输出全局决策,解码器部分从 pool 层输出局部决策。

此外,为增加模型的泛化能力,参考文献[21]中的方法在生成模型与判别模型中同时使用了 cutout 等正则化方法。

损失函数中原 StyleGAN2 网络的判别模型、生成模型的损失函数分别为

$$L_D = -E_x [\ln D(x)] - E_z \{ \ln \{ 1 - D[G(z)] \} \}, \quad (4)$$

$$L_G = -E_z \{ \ln D[G(z)] \}, \quad (5)$$

式中: $L_D$ 、 $L_G$  分别表示判别模型与生成模型的损失函数。改进后的判别模型损失函数为

$$L_D = L_{D_{enc}} + L_{D_{dec}}, \quad (6)$$

式中: $L_{D_{enc}}$ 、 $L_{D_{dec}}$  分别表示判别模型中编码器部分的损失函数与解码器部分的损失函数,可分别表示为

$$L_{D_{enc}} = -E_x \{ \ln D_{enc} [T(x)] \} - E_z \{ \ln \{ 1 - D_{enc} \{ T[G(z)] \} \} \}, \quad (7)$$

$$L_{D_{dec}} = -E_x \left[ \sum \ln \{ D_{dec} [T(x)] \} \right] - E_z \left\{ \sum \ln \{ 1 - \{ D_{dec} \{ T[G(z)] \} \} \} \right\}, \quad (8)$$

式中: $D_{enc}$ 、 $D_{dec}$  分别表示判别模型的编码器部分与

解码器部分;  $T$  表示 cutout 正则化方法;  $i, j$  分别表示判别模型解码部分  $D_{\text{dec}}$  输出矩阵的行、列数。改进后的生成模型损失函数为

$$L_G = -E_z \left\{ \ln D_{\text{enc}} \{ T [G(z)] \} + \sum \ln \{ D_{\text{dec}} \{ T [G(z)] \} \} \right\}. \quad (9)$$

### 3.3 生成目标图像融合

生成目标图像融合部分采用生成目标替换原边界框的方式,生成了与原始图像具有相同标签的异构图像,该融合过程主要分为纹理匹配与边界融合两步。其中纹理匹配是基于局部二值模式(LBP)<sup>[22]</sup>对生成图像与背景图像进行特征匹配,首先使用旋转不变 LBP 算子分别提取生成样本与背景图片的特征,之后计算二者的统计直方图特征,以生成样本与背景图片的统计直方图特征均方差作为评价指标,筛选与背景图片最为匹配的生成样本。生成样本与背景图片间的匹配损失的计算公式为

$$p_{\text{loss}} = (l_b - l_g)^2, \quad (10)$$

式中:  $l_b, l_g$  分别表示背景图片与生成样本的统计直方图特征。为保证生成图像的多样性,筛选过程中首先选择出最小的 10 张生成样本,再从 10 张生成样本中随机选择 1 张进行边界融合。

边界融合部分借鉴文献[23]中的边界渐进式融合的思想,设边界融合  $r$  个像素,融合时的权重  $w$  为

$$w = 1 - \frac{i}{r}, \quad i = 1, 2, \dots, r, \quad (11)$$

式中:  $i$  为边界融合的像素数,融合后图片边界处的像素为

$$g'(x, y) = w \cdot b(x, y) + (1 - w) \cdot g(x, y), \quad (12)$$

式中:  $g'(x, y)$  表示边界在  $(x, y)$  处融合后的像素;  $b(x, y)$  表示背景图片在  $(x, y)$  处的像素;  $g(x, y)$  表示生成图片在  $(x, y)$  处的像素。

## 4 实验与分析

### 4.1 数据集制作

本文选取 DIOR 数据集<sup>[24]</sup>与 RSOD 数据集<sup>[25]</sup>作为研究对象。DIOR 数据集共包含 23463 张遥感图像,其中训练集与验证集共 11725 张图像,测试集共 11738 张图像,图片分辨率为 800 pixel  $\times$  800 pixel,共 20 个类别。为研究数据增强算法在小样本数据集下对目标检测算法的影响,本文减少了训练集的数量,从训练集与验证集中随机抽取了 7035 张图像作为训练集,测试集保持不变。RSOD 数据集共包含 976 张遥感图像,图片分辨率为 512 pixel  $\times$  512 pixel 至 1961 pixel  $\times$  923 pixel 不等,图像共 4 个类别,本文按照 0.7, 0.1, 0.2 的比例划分了训练集、验证集与测试集。

### 4.2 实验结果和分析

本文分别选取了 DIOR 数据集中目标数量相对较少的 11 类目标与 RSOD 数据集中目标数量相对较少的 2 类目标进行逐类数据增强,表 1、2 分别为利用改进后的 StyleGAN2 与原 StyleGAN2 生成目标的 Frechet Inception 距离(FID 值)在两个数据集上的对比结果,实验结果为同等实验环境配置下模型的 FID 最小值。

表 1 DIOR 数据集上生成目标的 FID 值对比

Table 1 Comparison of FID values of generated target on DIOR datasets

Category	Airport	Basketball court	Bridge	Chimney	Overpass	Stadium
StyleGAN2	73.80	70.38	102.69	125.37	83.98	56.17
Ours	67.35	52.70	95.96	78.51	55.38	53.25
Category	Dam	Expressway-service-area	Golf field	Ground track field	Train station	
StyleGAN2	123.04	62.42	67.03	45.20	113.78	
Ours	103.00	54.22	57.98	39.87	93.00	

表 2 RSOD 数据集上生成目标的 FID 值对比

Table 2 Comparison of FID values of generated target on RSOD datasets

Category	Overpass	Playground
StyleGAN2	88.52	127.86
Ours	110.42	168.30

生成新的目标后,利用旋转不变 LBP 算子对生成目标与背景图片进行纹理匹配,筛选出与背景图片纹理最为接近的生成样本,并将其与背景图片融合后得到新的生成样本,生成的样本示例如图 4 所示。

在小样本数据集上分别使用基于图像仿射变换

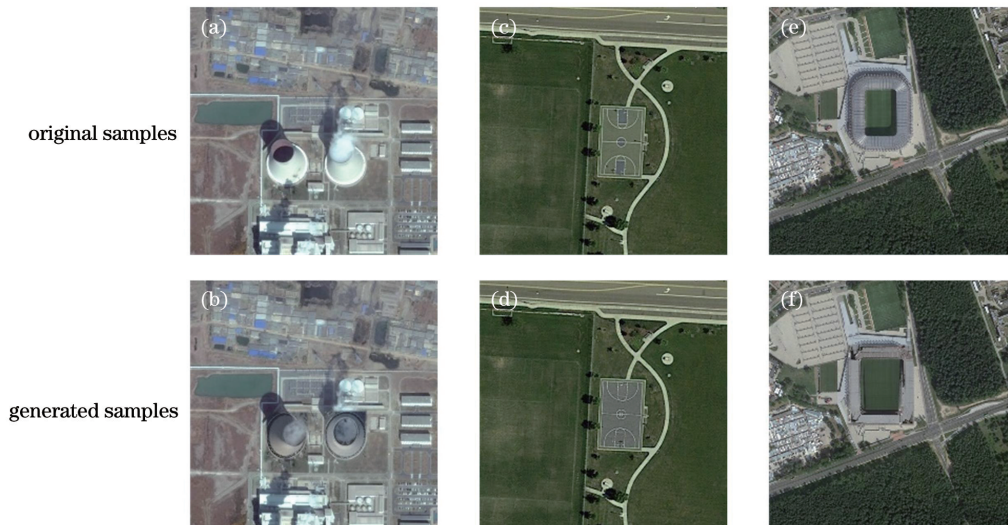


图 4 生成样本效果图示例。(a)(b)烟囱;(c)(d)篮球场;(e)(f)体育场

Fig. 4 Examples of generated samples. (a)(b) Chimney; (c)(d) basketball court; (e)(f) stadium

(包括平移变换与颜色变换)的方法以及本文所提方法进行数据增强,并通过在原数据集上增加不同数量

的增强数据进行对比,选择 yolov3 作为评价指标评价识别模型的检测精度,实验结果如表 3、4 所示。

表 3 DIOR 数据集上添加不同比例增强数据时的检测精度对比

Table 3 Comparison of detection accuracy for adding different percent of enhancement data on DIOR dataset %

Method	Percent of added data / Accuracy (mAP)				
Affine transformation	0/46.06	100/47.69	200/48.45	300/47.74	400/49.14
Ours	0/46.06	5/48.14	10/46.97	15/45.60	20/46.17
Combined method	0/46.06	105/48.09	205/49.48	305/47.96	405/49.51

表 4 RSOD 数据集上添加不同比例增强数据时的检测精度对比

Table 4 Comparison of detection accuracy for adding different percent of enhancement data on RSOD dataset %

Method	Percent of added data / Accuracy (mAP)				
Affine transformation	0/70.78	100/74.22	200/79.28	300/76.76	400/77.84
Ours	0/70.78	5/75.37	10/74.83	15/76.52	20/77.25
Combined method	0/70.78	120/79.34	220/80.13	320/77.77	420/78.35

从实验结果可以看出,基于图像仿射变换的方法在 DIOR 数据集与 RSOD 数据集上的检测精度分别为 46.06% 与 70.78%,最高精度分别为 49.14% 与 79.28%,使用本文方法进行数据增强后得到的最高精度分别为 48.14% 与 77.25%,略低于基于图像仿射变换的方法。但两种方法结合在一起时在 DIOR 数据集与 RSOD 数据集上取得的最高精度分别为 49.51% 与 80.13%,可见本文方法能对基于图像仿射变换的数据增强方法起到叠加的作用。

拟合的问题,针对少样本条件下的遥感图像数据,提出一种基于生成对抗网络的数据增强方法,通过实验验证了所提方法的有效性,该方法适用于遥感图像数据增强,且有助于提高检测识别模型的精度。

### 参 考 文 献

[1] Bjerrum E J. SMILES enumeration as data augmentation for neural network modeling of molecules [EB/OL]. (2017-05-17) [2020-07-30]. [https://www.researchgate.net/publication/315495821\\_SMILES\\_Enumeration\\_as\\_Data\\_Augmentation\\_for\\_Neural\\_Network\\_Modeling\\_of\\_](https://www.researchgate.net/publication/315495821_SMILES_Enumeration_as_Data_Augmentation_for_Neural_Network_Modeling_of_)

## 5 结 论

深度神经网络在小样本数据集下容易出现过

- Molecules.
- [2] Zhong Z, Zheng L, Kang G L, et al. Random erasing data augmentation [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 13001-13008.
- [3] Lemley J, Bazrafkan S, Corcoran P, et al. Smart augmentation learning an optimal data augmentation strategy [J]. IEEE Access, 2017, 5: 5858-5869.
- [4] Dinh L, Krueger D, Bengio Y, et al. NICE: non-linear independent components estimation [EB/OL]. (2014-10-30) [2015-04-10]. <http://arxiv.org/abs/1410.8516?context=cs.LG.html>.
- [5] Zhang X D, Wang T J, Yang Y, et al. Classification of small-sized sample hyperspectral images based on multi-scale residual network [J]. Laser & Optoelectronics Progress, 2020, 57(16): 162801. 张祥东, 王腾军, 杨耘, 等. 基于多尺度残差网络的小样本高光谱图像分类 [J]. 激光与光电子学进展, 2020, 57(16): 162801.
- [6] Jin W P, Guo J C, Qi Q, et al. Underwater image enhancement based on conditional generative adversarial network [J]. Laser & Optoelectronics Progress, 2020, 57(14): 141002. 晋玮佩, 郭继昌, 祁清, 等. 基于条件生成对抗网络的水下图像增强 [J]. 激光与光电子学进展, 2020, 57(14): 141002.
- [7] He Q, Li Y, Song W, et al. Multimodal remote sensing image classification with small sample size based on high-level feature fusion [J]. Laser & Optoelectronics Progress, 2019, 56(11): 111001. 贺琪, 李瑶, 宋巍, 等. 小样本的多模态遥感影像高层特征融合分类 [J]. 激光与光电子学进展, 2019, 56(11): 111001.
- [8] Cubuk E D, Zoph B, Mané D, et al. AutoAugment: learning augmentation strategies from data [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 113-123.
- [9] Terrance D, Graham W T. Improved regularization of convolutional neural networks with cutout [EB/OL]. (2017-08-15) [2017-11-29]. <http://arxiv.org/abs/1708.04552?context=cs.LG.html>.
- [10] Chen P, Liu S, Zhao H, et al. GridMask data augmentation [EB/OL]. (2020-01-13) [2020-01-14]. <http://arxiv.org/abs/2001.04086?context=cs.LG.html>.
- [11] Zhang H, Cisse M, Dauphin Y N, et al. Mixup: beyond empirical risk minimization [EB/OL]. (2017-10-25) [2018-04-27]. <http://arxiv.org/abs/1710.09412?context=cs.LG.html>.
- [12] Hendrycks D, Mu N, Cubuk E D, et al. AugMix: a simple data processing method to improve robustness and uncertainty [EB/OL]. (2019-11-05) [2020-02-17]. <http://arxiv.org/abs/1912.02781?context=cs.LG.html>.
- [13] Kingma D P, Welling M. Auto-encoding variational Bayes [EB/OL]. (2013-11-20) [2014-05-01]. <http://arxiv.org/abs/1312.6114?context=cs.LG.html>.
- [14] Goodfellow I, Pouget J, Mirza M, et al. Generative adversarial nets [C] // Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal, Quebec, Canada. New York: ACM, 2014: 2672-2680.
- [15] Karras T, Laine S, Aittala M, et al. Analyzing and improving the image quality of StyleGAN [EB/OL]. (2019-11-03) [2020-03-23]. <http://arxiv.org/abs/1912.04958?context=cs.LG.html>.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement [EB/OL]. (2018-04-08) [2020-03-23]. <http://arxiv.org/abs/1804.02767?context=cs.LG.html>.
- [17] Ronneberger O, Fischer P, Brox T, et al. U-net: convolutional networks for biomedical image segmentation [EB/OL]. (2015-05-18) [2020-03-23]. <https://arxiv.org/abs/1505.04597>.
- [18] Radford A, Metz L, Chintala S, et al. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2015-11-19) [2016-01-07]. <http://arxiv.org/abs/1511.06434?context=cs.LG.html>.
- [19] Arjovsky M, Chintala S, Bottou L. Wasserstein gan [EB/OL]. (2017-01-26) [2017-11-06]. <http://arxiv.org/abs/1701.07875?context=cs.LG.html>.
- [20] Karras T, Laine S, Aila T, et al. A style-based generator architecture for generative adversarial networks [EB/OL]. (2018-11-12) [2019-03-29]. <http://arxiv.org/abs/1812.04948?context=cs.LG.html>.
- [21] Zhao S, Liu Z, Lin J, et al. Differentiable augmentation for data-efficient gan training [EB/OL]. (2020-06-18) [2020-03-23]. <http://arxiv.org/abs/2006.10738?context=cs.LG.html>.
- [22] Ojala T, Pietikainen M, Maenpaa T, et al. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7): 971-987.
- [23] Li K, Wan G, Cheng G, et al. Object detection in optical remote sensing images: a survey and a new benchmark [EB/OL]. (2019-08-31) [2019-09-22].

<http://arxiv.org/abs/1909.00133?context=cs.LG.html>.

- [24] Liu Y Y, Zhang J M, Wang K P, et al. Fast underwater target recognition method based on unbalanced datasets[EB/OL]. (2019-07-19) [2019-07-30]. <http://kns.cnki.net/kcms/detail/11.2127>.

TP.20190719.1443.018.html.

- [25] Long Y, Gong Y P, Xiao Z F, et al. Accurate object localization in remote sensing images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2486-2498.