

基于改进生成对抗网络和 MobileNetV3 的带钢缺陷分类

常江¹, 管声启^{1,2*}, 师红宇³, 胡璐萍¹, 倪奕棋¹

¹ 西安工程大学机电工程学院, 陕西 西安 710048;

² 绍兴市柯桥区西纺纺织产业创新研究院, 浙江 绍兴 312030;

³ 西安工程大学计算机科学学院, 陕西 西安 710048

摘要 针对数据集样本数量较少会影响深度学习检测效果的问题, 提出了一种基于改进生成对抗网络和 MobileNetV3 的带钢缺陷分类方法。首先, 引入生成对抗网络并对生成器和判别器进行改进, 解决了类别错乱问题并实现了带钢缺陷数据集的扩充。然后, 对轻量级图像分类网络 MobileNetV3 进行改进。最后, 在扩充后的数据集上训练, 实现了带钢缺陷的分类。实验结果表明, 改进的生成对抗网络可生成比较真实的带钢缺陷图像, 同时解决深度学习中样本不足的问题; 且改进的 MobileNetV3 参数量是原有参数的 1/14 左右, 准确率为 94.67%, 比改进前提高了 2.62 个百分点, 可在工业现场对带钢缺陷进行实时准确的分类。

关键词 图像处理; 缺陷检测; 图像分类; 生成对抗网络; 数据增强

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0410016

Strip Defect Classification Based on Improved Generative Adversarial Networks and MobileNetV3

Chang Jiang¹, Guan Shengqi^{1,2*}, Shi Hongyu³, Hu Luping¹, Ni Yiqi¹

¹ School of Mechanical and Electronic Engineering, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China;

² Shaoxing Kegiao West-Textile Industry Innovative Institute, Shaoxing, Zhejiang 312030, China;

³ School of Computer Science, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China

Abstract Aiming at the problem that the small number of samples in the dataset will affect the effect of deep learning detection, a strip defect classification method based on improved generative adversarial networks and MobileNetV3 is proposed in this paper. First, a generative adversarial network is introduced, and the generator and discriminator are improved to solve the problem of category confusion and realize the expansion of the strip defect data set. Then, the lightweight image classification network MobileNetV3 is improved. Finally, it is trained on the expanded data set to realize the classification of strip defects. Experimental results show that the improved generative adversarial network can generate more real strip steel defect images and solve the problem of insufficient samples in deep learning. And the parameter amount of the improved MobileNetV3 is about 1/14 of that before improvement, and the accuracy is 94.67%, which is 2.62 percentage points higher than that before improvement. It can be used for accurate and real-time classification of strip steel defects in industrial field.

Key words image processing; defect detection; image classification; generative adversarial networks; data augmentation

OCIS codes 100.2000; 120.4290; 100.4996; 150.1835

收稿日期: 2020-07-06; 修回日期: 2020-07-16; 录用日期: 2020-08-06

基金项目: 陕西省重点研发计划(2018GY-020)、绍兴市柯桥区西纺纺织产业创新研究院(19KQYB13)

* E-mail: sina1300841@163.com

1 引言

在汽车、船舶、航空航天等工业领域,带钢是重要的原材料之一,其质量直接影响了工业产品的最终性能。在带钢生产过程中,由于加工工艺、轧制设备不同,会出现孔洞、划痕、裂纹等各种各样的缺陷^[1-2]。人工检测带钢缺陷效率不高且容易受主观因素的影响。近年来,人们通过引入深度学习进行视觉检测^[3-8]。深度学习是机器学习中一种对数据进行表征学习的方法^[9],可用于带钢缺陷的自动检测。Vannocci 等^[10]分别用传统机器学习和深度学习对热轧带钢缺陷进行分类,发现深度学习的性能更好。王立中等^[11]建立卷积神经网络(CNN)并通过参数优化实现了有缺陷和无缺陷带钢的准确分类。Mi 等^[12]对密集连接网络(DenseNet)进行改进并将其用于钢丝表面的缺陷分类,准确率为 97.2%。基于大数据训练的深度学习算法能有效解决缺陷检测中出现的问题,但受工业现场数据采集条件的限制,采集的数据集样本数量较少。因此,还需进一步研究如何扩充小样本数据集^[13]。

Goodfellow 等^[14]提出生成对抗网络(GAN)模型,包括生成器和判别器两部分,二者进行对抗训练,最终可根据原始数据的特点生成逼真的样例。Mirza 等^[15]提出条件 GAN(CGAN),使 GAN 可同时生成多种类别的图像。Redford 等^[16]将 CNN 引入 GAN,提出深度卷积 GAN(DCGAN),为生成更高分辨率的图像奠定了基础。近年来,人们用 GAN 扩充数据集并取得了较好的效果。Xuan 等^[17]采用 GAN 生成多视角珍珠图像,并在扩充后的数据集上训练 CNN,显著降低了珍珠分类的错误率。Yi 等^[18]用 DCGAN 扩充后的数据集训练行人检测模型,准确率可达到 95% 以上。本文引入 GAN 并从生成器和判别器两方面进行改进,以生成更真实的带钢缺陷图像。然后,用改进后的轻量级神经网络 MobileNetV3 在扩充后的数据集上训练,最终实现带钢缺陷分类。

2 基本原理

2.1 改进的生成对抗网络

辅助分类器 GAN(AC-GAN)^[19]是一种将 CGAN 和 DCGAN 结合的多类图像生成模型。AC-GAN 的目标函数由正确来源的对数似然估计 L_S 和正确类别的对数似然估计 L_C 组成。判别器的训练目标是最大化 $L_C + L_S$,生成器的训练目标是最大

化 $L_C - L_S$ 。正确来源的对数似然估计 L_S 和正确类别的对数似然估计 L_C 可表示为

$$L_S = E[\log P(S = r | X_{\text{real}})] + E[\log P(S = f | X_{\text{fake}})], \quad (1)$$

式中, X_{real} 为真实图像, X_{fake} 为生成器生成的图像, S 为图像来源, $S = r$ 表示图像来源于真实图像, $S = f$ 表示图像来源于生成器生成的图像。

$$L_C = E[\log P(C = c | X_{\text{real}})] + E[\log P(C = c | X_{\text{fake}})], \quad (2)$$

式中, C 为判别器得到的类别, c 为真实类别。

原始 AC-GAN 通过 Embedding 稠密编码将图像标签映射到和随机噪声相同的维度下,然后将二者相乘并输入生成器,经过逐层转置卷积处理,图像标签的影响力会逐渐衰减。原始 AC-GAN 将判别器的最后一层卷积特征图展平之后进行分支,分别全连接到判别真假和判别类别的输出层。但该方法存在的缺点:1)卷积层到全连接层需要将二维特征图展成一维,大幅度增加了全连接层和输出层的节点数,不利于模型训练;2)判别真假和判别类别的支路都只有一层隐含层,限制了支路的拟合能力。针对 AC-GAN 的缺点,实验从生成器和判别器两方面进行改进。

改进的 AC-GAN 生成器网络结构如图 1 所示,其中, noises 为高斯白噪声,是一组服从标准正态分布的随机数。Transposed Conv 为卷积核尺寸为 4×4 的转置卷积操作。 $4 \times 4 \times 512$ 为一组特征图的尺寸, 4×4 为特征图的宽和高, 512 为特征图的通道数。DCGAN 可将深度卷积引入 GAN 并通过实验确定卷积通道数等参数,因此,实验中生成器和判别器的卷积特征图尺寸及通道数设置参考 DCGAN。

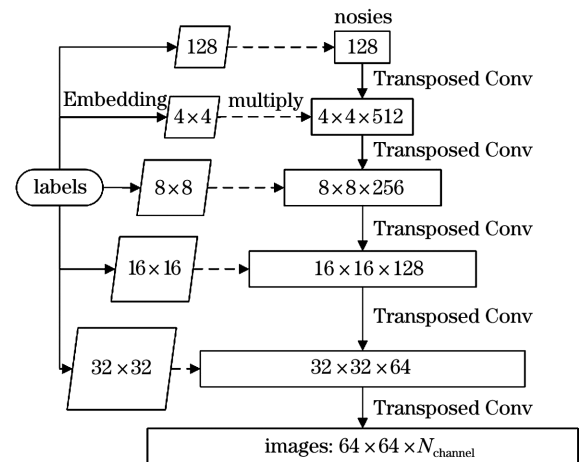


图 1 生成器网络的结构

Fig. 1 Structure of the generator network

实验使用的带钢缺陷图像为灰度图,即图像通道数 N_{channel} 为 1。将图像标签经过稠密编码 Embedding 分别映射到 $128, 4 \times 4, 8 \times 8, 16 \times 16, 32 \times 32$ 等维度,然后和生成器每层网络对应尺寸的特征图相乘后再进行转置卷积,从而将图像标签的影响力贯穿整个生成器网络。

改进的 AC-GAN 判别器网络结构如图 2 所示,其中,Conv 4×4 表示卷积核尺寸为 4×4 的卷积操作。在判别器第三次卷积得到尺寸为 $8 \times 8 \times 256$ 的特征图后进行分支,各支路分别进行三次卷积得到判别结果。 1×1 卷积可在不改变特征图尺寸的前提下增加网络深度,且不会引入太多计算量。因此,实验引入 1×1 卷积给各支路增加一层隐含层。改进的判别器没有全连接层,避免了节点数大幅度增加和减少的异常,但各支路的隐含层数增多,既能保证支路之间的权值共享,又增强了各自的独立拟合能力。

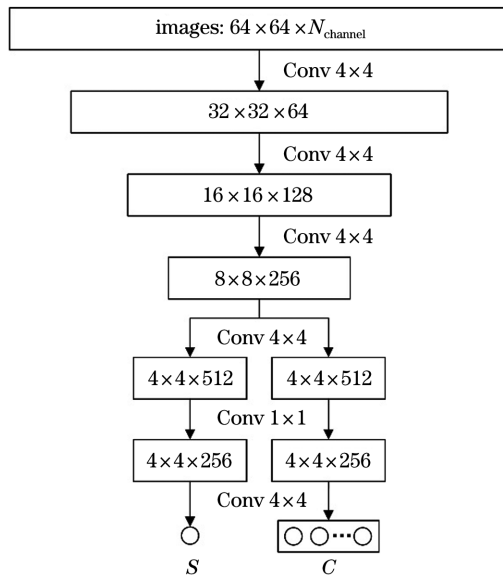


图 2 判别器网络的结构

Fig. 2 Structure of the discriminator network

2.2 改进的 MobileNetV3

实验使用深度学习中的图像分类网络实现带钢缺陷分类。一些经典的图像分类网络为了达到较高的准确率会增加网络结构的复杂度,导致参数量大、占用内存多、预测速度慢。近年来,人们开始研究可在手机等性能受限设备上运行的轻量级神经网络。MobileNetV3^[20] 在训练中采用互补搜索技术搜寻轻量高效的网络架构,图像分类性能较高,且参数量少,预测速度快,可满足工业应用的要求。实验使用的带钢缺陷图像复杂程度较低,识别难度也较低,因此,对 MobileNetV3 网络架构进行改进,调整卷积

层和通道数,在减少网络数量的同时提高运算速度,使网络更适用于带钢缺陷分类。

改进后的 MobileNetV3 网络结构如表 1 所示,其中,Resolution 为图像的分辨率,Channel 为输出通道数,Layer 为网络层数。改进后的网络在保留关键模块的前提下去除原始 MobileNetV3 的一些冗余结构,如将表 1 倒数第二个卷积层的通道数的 Channel 由 1280 改为 240,将第三个 Bottleneck 5×5 的网络层数由 5 改为 2。单纯减少卷积层和通道数可能会使模型拟合能力下降,因此,将除输出层外所有层的激活函数都设为 Hard Swish (原始 MobileNetV3 部分层的激活函数为线性修正单元 ReLU,非线性不如 Hard Swish)。

表 1 改进 MobileNetV3 的结构

Table 1 Structure of the improved MobileNetV3

Operator	Resolution	Channel	Layer
Conv 3×3	64×64	16	1
Bottleneck 3×3	32×32	24	2
Bottleneck 5×5	16×16	40	2
Bottleneck 5×5	8×8	80	2
Conv 1×1 & Pooling	4×4	120	1
Conv 1×1	1×1	240	1
Conv 1×1	1×1	6	1

3 实验结果与分析

所有实验均在同一台服务器上完成,CPU 为酷睿 i7-7820X,GPU 为 2 个 GeForce RTX2080Ti,运行内存 32 G。采用 Song 等^[21]制作的 NEU 带钢表面缺陷数据集,包含 6 类缺陷,每类缺陷包含 300 张灰度图像(尺寸均为 $64 \text{ pixel} \times 64 \text{ pixel}$)。图 3 为部分原始带钢缺陷图像,第一列为细纹缺陷,第二列为杂质缺陷,第三列为斑块缺陷,第四列为点蚀缺陷,第五列为轧入氧化皮缺陷,第六列为划痕缺陷。

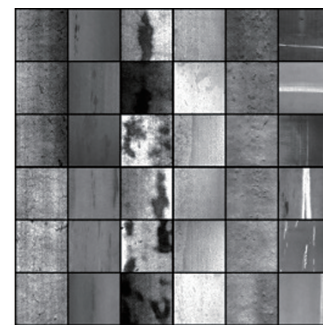


图 3 部分原始带钢缺陷图像

Fig. 3 Part of the original strip steel defect images

为了验证改进的 AC-GAN 生成图像的有效性, 分别将改进前后的 AC-GAN 在原始带钢缺陷数据集上训练 500 轮。图 4(a1)~4(c1)、图 4(a2)~图 4(c2)分别为原始 AC-GAN 和改进 AC-GAN 训练 100 轮、300 轮、500 轮生成的图像。理想情况下, 单个图像生成效果图从第一列到第六列应该分别为细纹、杂质、斑块、点蚀、轧入氧化皮、划痕六种不同

的缺陷。但 AC-GAN 在图 4(a1)第二列杂质缺陷第二行、图 4(b1)第二列杂质缺陷最后一行等位置出现了类别错乱的情况。而改进的 AC-GAN 类别清晰, 从第 100 轮到第 500 轮, 图像生成效果越来越好, 到第 500 轮模型已经收敛, 和图 1 的原始图像非常接近, 这表明改进的 AC-GAN 适用于带钢缺陷数据集扩充。

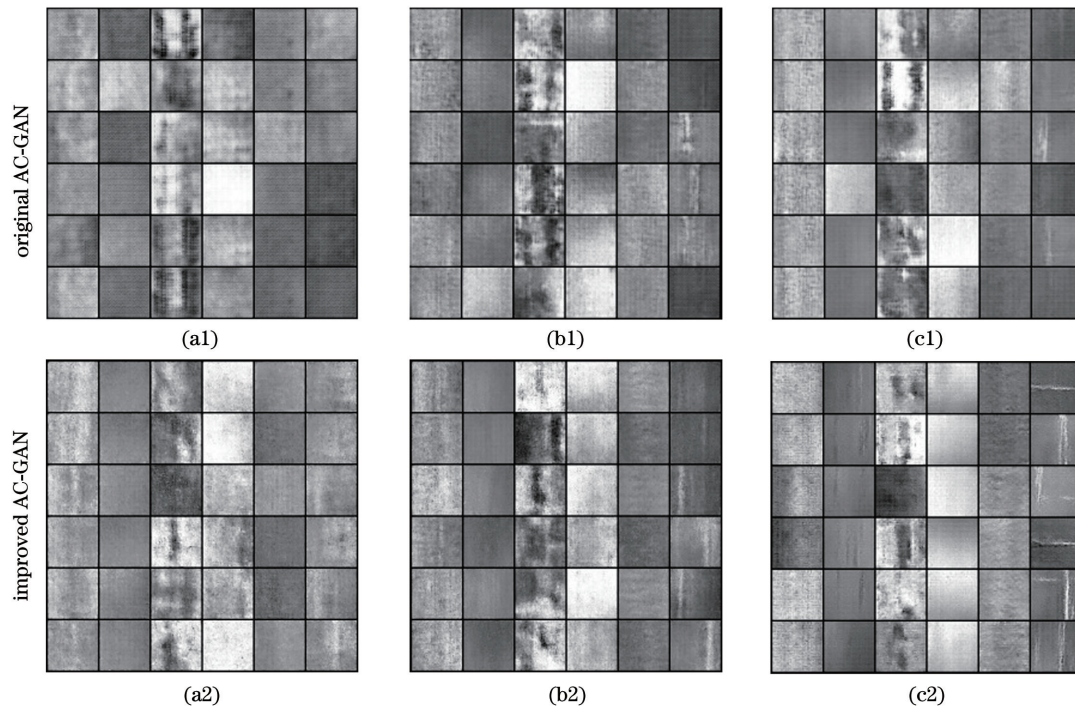


图 4 不同网络训练得到的图像。(a)100 轮;(b)300 轮;(c)500 轮

Fig. 4 Images obtained by different network training. (a) 100 epochs; (b) 300 epochs; (c) 500 epochs

将原始带钢缺陷数据集分为训练集和测试集两部分, 训练集每类缺陷包含 200 张图像, 测试集每类包含 100 张图像。然后, 用训练好的改进 AC-GAN 将训练集每种缺陷的 200 张图像扩充到 1200 张。为了验证改进的 MobileNetV3 在带钢缺陷分类中的有效性, 分别用原始 MobileNetV3、改进的 MobileNetV3、VGG16、ResNet34、DenseNet121 在扩充后的带钢缺陷数据

集上训练 100 轮。由于带钢缺陷图像的认识难度较小, 因此没有选用复杂的网络作为对比, 实验结果如表 2 所示。其中, Parameter amount 为网络的卷积核、权重等参数数量, 单位为百万。Predicted time 为预测完测试集 600 张图像所用的平均时间, Accuracy(准确率)、Precision(精确率)、Recall(召回率)和 F_1 分数都是测试集中每类图像对应指标的平均值。

表 2 不同网络的图像分类性能

Table 2 Image classification performance of different networks

Network	Parameter amount /million	Predicted time /ms	Accuracy /%	Precision /%	Recall /%	F_1
Original MobileNetV3	1.24	272	92.05	92.54	91.47	0.92
Improved MobileNetV3	0.09	231	94.67	95.21	94.79	0.95
VGG16	17.35	320	93.85	94.33	93.67	0.94
ResNet34	21.28	318	91.23	91.41	90.59	0.91
DenseNet121	6.95	315	94.36	94.50	94.10	0.94

从表 2 可以发现,改进 MobileNetV3 的参数量是原始 MobileNetV3 的 1/14 左右,分别为 VGG16、ResNet34、DenseNet121 的 1/200、1/237、1/77 左右。改进 MobileNetV3 的预测时间比原始 MobileNetV3 快 1/6 左右,比其他网络快 2/7 左右。神经网络的预测时间受运算操作、激活函数等多种因素影响且与参数量不是线性相关的,但参数量下降后预测时间也有明显缩短,这有利于实时检测。改进 MobileNetV3 的准确率比原始 MobileNetV3 高 2.62 个百分点,分别比 VGG16、ResNet34、DenseNet121 高 0.82、3.44、0.31 个百分点。 F_1 分数将精确率和召回率结合,可评价图像分类的综合性能。改进 MobileNetV3 的 F_1 分数比原始 MobileNetV3 高 0.03,分别比 VGG16、ResNet34、DenseNet121 高 0.01、0.04、0.01。这表明改进 MobileNetV3 的各项图像分类指标均优于其他网络,原因是带钢缺陷图像的复杂程度较低,神经网络结构太复杂效果反而不好,实验使用轻量级神经网络 MobileNetV3 并将其进一步精简,因此可以准确高效地分类带钢缺陷。

4 结 论

为了扩充带钢缺陷数据集,首先,引入生成对抗网络 AC-GAN 并对其改进,将图像标签映射到不同维度并与生成器紧密融合;然后,通过增加卷积层提高判别器中判别真假和判别类别两条支路的拟合能力。最后,基于改进 AC-GAN 扩充后的数据集对改进的 MobileNetV3 进行训练,实现带钢缺陷的高效分类。实验结果表明,改进的 AC-GAN 可以很好地解决原始 AC-GAN 的类别错乱问题且生成比较真实的带钢缺陷图像。同时,改进 MobileNetV3 的参数量和预测时间有大幅降低,图像分类准确率等各项指标也优于实验中的其他网络,可对带钢缺陷进行准确实时分类。目前只验证了改进的 AC-GAN 生成尺寸为 64 pixel \times 64 pixel 图像的有效性,之后还需进一步研究更高分辨率图像的生成。

参 考 文 献

- [1] Ghorai S, Mukherjee A, Gangadaran M, et al. Automatic defect detection on hot-rolled flat steel products[J]. IEEE Transactions on Instrumentation and Measurement, 2013, 62(3): 612-621.
- [2] Neogi N, Mohanta D K, Dutta P K. Review of vision-based steel surface inspection systems [J]. EURASIP Journal on Image and Video Processing, 2014, 2014(1): 1-19.
- [3] Saito K, Ushiku Y, Harada T, et al. Strong-weak distribution alignment for adaptive object detection [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: 6949-6958.
- [4] Yang F, Fan H, Chu P, et al. Clustered object detection in aerial images [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 8310-8319.
- [5] Yan Z Y, Yuan Y C, Zuo W M, et al. Perspective-guided convolution networks for crowd counting[C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 952-961.
- [6] Shen X H, Li Z H, Li M, et al. Aluminum surface-defect detection based on multi-task deep learning[J]. Laser & Optoelectronics Progress, 2020, 57(10): 101501.
沈晓海, 栗泽昊, 李敏, 等. 基于多任务深度学习的铝材表面缺陷检测[J]. 激光与光电子学进展, 2020, 57(10): 101501.
- [7] Zhang G S, Ge G Y, Zhu R H, et al. Gear defect detection based on the improved YOLOv3 network [J]. Laser & Optoelectronics Progress, 2020, 57(12): 121009.
张广世, 葛广英, 朱荣华, 等. 基于改进 YOLOv3 网络的齿轮缺陷检测[J]. 激光与光电子学进展, 2020, 57(12): 121009.
- [8] Liu F, Wu Z W, Yang A Z, et al. Multi-scale feature fusion based adaptive object detection for UAV [J]. Acta Optica Sinica, 2020, 40(10): 1015002.
刘芳, 吴志威, 杨安喆, 等. 基于多尺度特征融合的自适应无人机目标检测[J]. 光学学报, 2020, 40(10): 1015002.
- [9] Bengio Y, Delalleau O. On the expressive power of deep architectures [M] // Elomaa T, Hollmén J, Mannila H, et al. Discovery Science. DS 2011. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2011, 6926: 1.
- [10] Vannocci M, Ritacco A, Castellano A, et al. Flatness defect detection and classification in hot rolled steel strips using convolutional neural networks [C] // 15th International Work Conference on Artificial Neural Networks (IWANN), June 12-14, 2019, Gran Canaria, Spain. Cham: Springer, 2019: 220-234.
- [11] Wang L Z, Guan S Q. Strip steel surface defect

- recognition based on deep learning [J]. Journal of Xi'an Polytechnic University, 2017, 31(5): 669-674.
- 王立中, 管声启. 基于深度学习算法的带钢表面缺陷识别[J]. 西安工程大学学报, 2017, 31(5): 669-674.
- [12] Mi Z S, Song Y H, Yan Y. A defect classification network based on deformation dense connection in wire rod surface image [C] // 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI), September 21-22, 2019, Xi'an, China. New York: IEEE Press, 2019: 155-160.
- [13] Liu S Y, Guo H Y, Hu J G, et al. A novel data augmentation scheme for pedestrian detection with attribute preserving GAN [J]. Neurocomputing, 2020, 401: 123-132.
- [14] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C] // 27th International Conference on Neural Information Processing Systems(NIPS), December 8-13, 2014, Montreal, Canada. Cambridge: MIT Press, 2014: 2672-2680.
- [15] Mirza M, Osindero S. Conditional generative adversarial nets [EB/OL]. [2020-06-20]. <https://arxiv.org/abs/1411.1784>.
- [16] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. [2020-06-24]. <https://arxiv.org/abs/1511.06434>.
- [17] Xuan Q, Chen Z Z, Liu Y, et al. Multiview generative adversarial network and its application in pearl classification [J]. IEEE Transactions on Industrial Electronics, 2019, 66(10): 8244-8252.
- [18] Yi C, Cho J. Improving the performance of multimedia pedestrian classification with images synthesized using a deep convolutional generative adversarial network [J]. Multimedia Tools and Applications, 2020, 89: 1-16.
- [19] Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs [EB/OL]. [2020-06-21]. <https://arxiv.org/abs/1610.09585>.
- [20] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 1314-1324.
- [21] Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects [J]. Applied Surface Science, 2013, 285: 858-864.