



# 基于语义分割的桥梁锈蚀病害 识别对比分析

倪有豪 陆欢 季超 茅建校 王浩 徐寅飞

(东南大学混凝土及预应力混凝土结构教育部重点实验室, 南京 211189)

**摘要:** 针对桥梁锈蚀数据难获取、锈蚀病害数据集小的问题,基于生成对抗网络(GAN)对桥梁锈蚀数据集进行扩增,并采用 IS 和 K 均值聚类算法验证其有效性.采用扩增后的数据集,按 4:1 的比例划分为训练集和验证集,分别对 U-Net 网络和 DeepLab-V3+ 网络进行训练,对比分析 2 种网络对于锈蚀分割的精度、召回率及  $F_1$  分数.结果表明,采用深度卷积生成对抗网络(DCGAN)生成虚拟数据集的 IS 值达到 2.41,分类肘形图类别数为 5,与原数据集吻合,可作为扩增数据集以提升模型泛化性;DeepLab-V3+ 网络模型在验证集上的精度为 0.935,召回率为 0.952,  $F_1$  分数为 0.943,均显著高于 U-Net 网络模型. DeepLab-V3+ 网络在点状锈蚀区域连通与分割方面优于 U-Net 网络,并实现了片状锈蚀区域分割,为桥梁锈蚀精准识别与分割提供了技术支撑.

**关键词:** 深度学习;生成对抗网络;桥梁锈蚀识别;语义分割

**中图分类号:** U446.3 **文献标志码:** A **文章编号:** 1001-0505(2023)02-0201-09

## Comparative analysis on bridge corrosion damage detection based on semantic segmentation

Ni Youhao Lu Huan Ji Chao Mao Jianxiao Wang Hao Xu Yinfei

(Key Laboratory of Concrete and Prestressed Concrete Structures of Ministry of Education, Southeast University, Nanjing 211189, China)

**Abstract:** To solve the problems that the data of bridge corrosion is difficult to obtain and the dataset of corrosion damage is small, the bridge corrosion dataset was amplified based on generative adversarial networks (GAN). Inception score (IS) and K-means clustering algorithm were adopted to verify the effectiveness of the amplified dataset. The amplified dataset was divided into a training dataset and a validation dataset at a ratio of 4:1. U-Net network and DeepLab-V3+ network were respectively trained by the training dataset. The precision, the recall and  $F_1$  score of the two networks for corrosion segmentation were compared. The results show that the IS value of the synthetic dataset generated by the deep convolutional generative adversarial networks (DCGAN) is 2.41, and the category number of the classification elbow digram is 5, which is consistent with the original dataset. The synthetic dataset can be added as amplified datasets to improve model generalization. The precision, the recall rate, and the  $F_1$  score of the DeepLab-V3+ network model are 0.935, 0.952, and 0.943 in the validation dataset, respectively, which are significantly higher than those of the U-Net network model. DeepLab-V3+ network is superior to U-Net network in connection and segmentation of spot-corrosion area, and realizes segmentation of flake-corrosion area, providing technical support for accurate detection and segmentation of bridge corrosion.

**Key words:** deep learning; generative adversarial networks; corrosion detection of bridges; semantic segmentation

中国桥梁建设进入建养并重期,随着桥梁年龄的增加和交通压力的增大,桥梁病害日益增多.为

消除桥梁病害可能带来的安全隐患,针对在役桥梁的病害检测势在必行.桥梁锈蚀作为典型的桥梁病

**收稿日期:** 2022-06-29. **作者简介:** 倪有豪(1987—),男,博士生;茅建校(联系人),男,博士,副研究员, jianxiao@seu.edu.cn.

**基金项目:** 国家自然科学基金资助项目(51978155,52108274)、江苏省研究生培养创新工程研究生科研与实践创新计划资助项目(SJCX21\_0056).

**引用本文:** 倪有豪,陆欢,季超,等.基于语义分割的桥梁锈蚀病害识别对比分析[J].东南大学学报(自然科学版),2023,53(2):201-209.

DOI: 10.3969/j.issn.1001-0505.2023.02.003.

害,会降低桥梁承载力,缩短桥梁使用寿命,造成巨大的经济损失<sup>[1]</sup>. 由于锈蚀形状不规则、难以辨认,针对桥梁锈蚀的检测,目前主要依靠人工目视. 这种传统的检测方法实时性差,检查范围不全面,影响交通通行,且成本过高<sup>[2]</sup>,无法满足巡检需求.

近年来,计算机视觉方法的出现为解决锈蚀识别提供了新思路. Petricca 等<sup>[3]</sup>通过对比传统计算机视觉技术和深度学习方法在锈蚀检测上的性能,发现深度学习分类器模型在精度和检出率方面明显优于基于 OpenCV 的颜色空间阈值处理方法. Deegan 等<sup>[4]</sup>对锈蚀图像进行区域分类,划分为完全锈蚀和完全没有锈蚀 2 个类别,构建 CNN 模型,利用滑动窗口算法进行训练和测试,得到了较高的准确率,但对于包含部分锈蚀的图像无法实现判断. 王达磊等<sup>[5]</sup>基于 U-Net 网络,利用苏通大桥的实景锈蚀图像进行训练和测试,实现了锈蚀面积的定量检测,但锈蚀的图像无法判断锈蚀程度. Chen 等<sup>[6]</sup>进一步将锈蚀划分为无锈蚀、轻微、中度、重度 4 个程度进行分类和锈蚀等级评估,展开了不同网络模型的对比实验. Bastian 等<sup>[7]</sup>利用滑动窗口算法,将识别腐蚀的亚区域映射至原图像中,并采用区域递归方法实现中度和重度锈蚀区域的框选展示,但并没有给出较为精确的锈蚀边界. Qian<sup>[8]</sup>基于 U-Net、DeepLab、PSPNet、RefineNet 四种网络展开对比实验,在锈蚀图像分割的应用上给出了不同网络的优缺点,但没有针对大面积、边缘模糊的锈蚀进行训练和测试. Xu 等<sup>[9]</sup>提出基于 U-Net 和 DeepLab-V3 的 ECNN 集成学习模型,在锈蚀率和锈蚀等级的识别精度上都高于单一分类器,但其利用的锈蚀图像经过裁剪后边界轮廓趋于平滑,没有完整体现锈蚀图像的特征.

针对桥梁锈蚀数据集锈蚀种类少、数据获取难且样本总量小的问题,本文采用生成对抗网络 (generative adversarial networks, GAN) 进行数据集扩增<sup>[10]</sup>,并利用 IS 和 K 均值聚类算法进行验证. 然后,基于 DeepLab-V3 + 网络对锈蚀数据集展开训练和验证,并与 U-Net 网络的锈蚀语义分割性能进行对比分析.

## 1 基于 GAN 的数据集扩增

### 1.1 GAN 原理

GAN<sup>[11]</sup>基本原理是由 1 个生成器和 1 个判别器形成互相博弈机制,即

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

<http://journal.seu.edu.cn>

式中,  $x$  表示真实图像;  $p_{\text{data}}(x)$  表示真实图像集服从的分布;  $D(x)$  表示判别  $x$  为真实图像的概率;  $p_z(z)$  表示随机噪声  $z$  服从的概率分布;  $G(z)$  表示生成器,输出值为生成图像;  $D(G(z))$  为判别器判断生成器生成图像为真实图像的概率;  $E_{x \sim p_{\text{data}}(x)} [\log D(x)]$  表示输入真实图像时判别器能判别的期望值;  $E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$  表示输入生成器生成图像时判别器能判别的期望值;  $V(D, G)$  表示目标损失函数.

最优生成器  $G^*$  应满足

$$G^* = \arg \min_D \max_D V(D, G) \quad (2)$$

式中,  $\max_D V(D, G)$  表示使  $V(D, G)$  值最大化的判别器;  $\arg \min_D \max_D V(D, G)$  表示当判别器固定时,使  $V(D, G)$  值最小化的生成器.

训练判别器时,为区分真实图像与生成器生成的虚拟图像,  $D$  应使  $V(D, G)$  取值最大. 对于服从真实图像分布的  $x$ ,  $D(x)$  应接近于 1; 对于服从生成虚拟图像分布的  $x$ ,  $D(x)$  应接近于 0. 训练生成器时,固定判别器,应使  $V(D, G)$  取值最小,生成真实图像以欺骗判别器. 对于服从生成虚拟图像分布的  $x$ ,  $D(x)$  应接近于 1. 当判别器对于任意测试图像的正确识别率为 0.5 时,判别器无法区分真实图像与生成的虚拟图像.

本文选取 GAN 的变体,即深度卷积生成对抗网络 (deep convolutional generative adversarial networks, DCGAN)<sup>[12]</sup>. 该网络在生成器和判别器中使用卷积神经网络替代多层感知机,并采用卷积层取代全连接层,将上采样改为带步长的卷积,从而更好地提取图像特征. ReLU 激活函数在输入值小于 0 的区间上梯度为 0,而 Leaky ReLU 激活函数在输入值小于 0 的区间上梯度为 0.1,保持梯度稳定. 因此,DCGAN 的判别器激活函数采用 Leaky ReLU 替代 ReLU,以防止梯度消失. 生成器激活函数仍然采用 ReLU,但输出层采用 tanh 函数. 判别器和生成器的每一层均采用批量归一化 (BN) 层,以增加 DCGAN 训练的效率和稳定性.

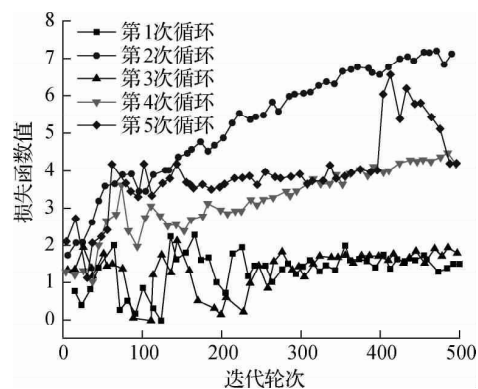
### 1.2 网络训练过程

DCGAN 的训练流程如下:

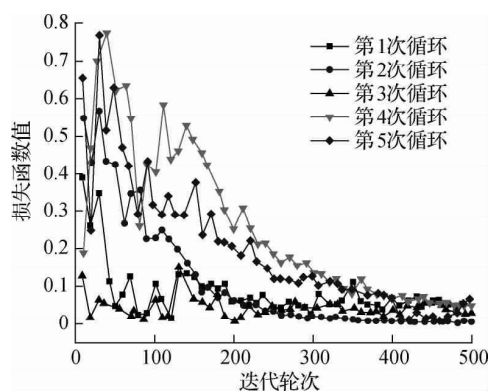
- ① 初始化生成器和判别器的网络参数.
- ② 从训练集抽取单个样本,生成器利用定义的噪声分布生成单个样本. 固定生成器,训练判别器,使其尽可能区分真实图像与虚拟图像.
- ③ 循环更新  $k$  次判别器后,更新 1 次生成器,使判别器尽可能无法区分真实图像与虚拟图像.

通过现场拍摄桥梁和网络爬虫方式获取了

1 200张桥梁锈蚀图像,包含混凝土桥梁剥落及露筋锈蚀图像、钢结构桥梁锈蚀图像.利用训练后得到的生成器生成图像.训练中的生成器和判别器的损失函数值变化曲线见图1.由图可知,生成器损失函数在震荡后趋于稳定,判别器损失函数先上升后迅速下降并趋于平稳.图2为桥梁锈蚀真实图像;图3为生成器生成的桥梁锈蚀图像.

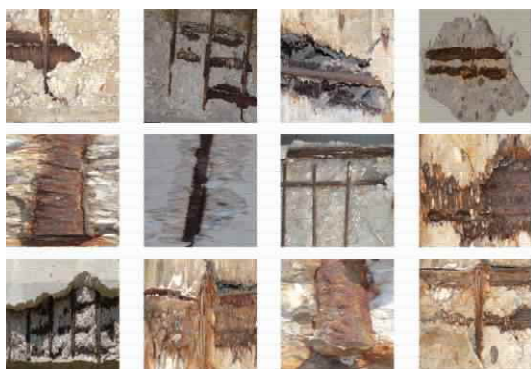


(a) 生成器



(b) 判别器

图1 DCGAN训练过程中损失函数值变化曲线

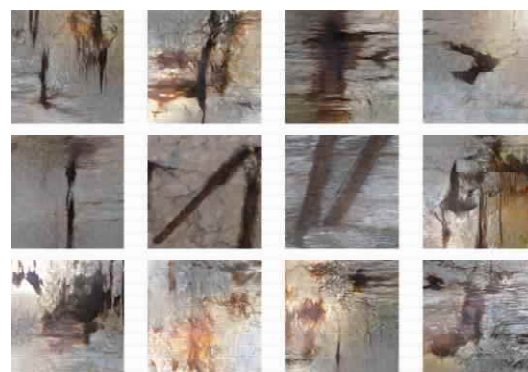


(a) 混凝土锈蚀



(b) 钢结构锈蚀

图2 桥梁锈蚀真实图像



(a) 混凝土锈蚀



(b) 钢结构锈蚀

图3 生成器生成的桥梁锈蚀图像

由图2(a)可见,混凝土锈蚀图像一般为混凝土剥落后露筋锈蚀,呈现为铁褐色或棕黄色,光泽暗淡,边缘多为不规则形状且较模糊,常呈裂齿状.当锈蚀发展到一定程度,会连成片状的锈蚀.由图2(b)可见,钢结构表面锈蚀一般早期为点状锈蚀,逐渐发展为块状锈蚀,最终为片状锈蚀,颜色多为棕色.结合图2和图3可以发现,生成的虚拟锈蚀图像在颜色、光泽、锈迹形态及锈蚀边缘形状方面与真实锈蚀图像吻合度较高,在钢筋锈蚀纹理方面存在些许差别.同时,虚拟锈蚀图像的分辨率较低.

## 2 DCGAN生成数据集评估

针对GAN的量化评估指标包括IS<sup>[13]</sup>和距离评估指标FID<sup>[14]</sup>.考虑到FID是衡量多元正态分布之间的距离,而神经网络提取的特征并不是多元正态分布的,故本文选用IS作为定量分析的主要评估指标.IS忽略了数据集之间的类别联系,因此本文同时采用K均值聚类算法<sup>[15]</sup>比较原始数据集和生成数据集的聚类数.

### 2.1 IS方法评估分析

IS从如下2个方面评估生成器的质量:

1) 批量生成图像,得到批量输出分布的信息熵,再求熵的期望,其定义式为

$$E_{x_G \sim P_G}(H(p(y|x_G))) = \sum_{x_G \in G} P(x_G) \sum_{i=1}^n P(y_i|x_G) \log \frac{1}{P(y_i|x_G)} \quad (3)$$

式中, $H(\cdot)$ 表示随机变量的信息熵; $E_{x_G \sim P_G}(\cdot)$ 表示在生成器生成图像 $x_G$ 服从概率分布 $P_G$ 时的随机变量期望值; $y$ 表示预训练中Inception所预测

<http://journal.seu.edu.cn>

的图像标签;  $P(x_G)$  表示生成器  $G$  生成图像  $x_G$  的概率;  $p(y|x_G)$  表示 Inception 输入图像为  $x_G$  时的输出类别分布;  $P(y_i|x_G)$  表示 Inception 预测输入图像  $x_G$  属于  $y_i$  标签类别的概率, 且  $i \in \{1, 2, \dots, n\}$ ,  $n$  为生成图像的类别总数.

2) 批量生成图像, 得到批量输出分布的期望, 再求期望的信息熵, 其定义式为

$$H(E_{x_G \sim P_G}(p(y|x_G))) = -\sum_{x_G \in G} P(x_G) \sum_{i=1}^n P(y_i|x_G) \log \frac{1}{P(y_i)} \quad (4)$$

综合式 (3) 和 (4), 可以得到 IS 的定义式为

$$Q_{IS} = \exp(E_{x_G \sim P_G} D_{KL}(p(y|x_G) \| p(y))) \quad (5)$$

式中,  $D_{KL}$  为条件分布  $p(y|x_G)$  与类别标签的边缘分布  $p(y)$  间的 KL 散度, 衡量 2 个分布之间的距离, 且

$$D_{KL}(P \| Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \quad (6)$$

式中,  $P(i)$ 、 $Q(i)$  均为概率分布函数.

$D_{KL}$  值越大,  $P(i)$ 、 $Q(i)$  两个概率分布的差异性越大.  $P(y|x_G)$  的熵值大小反映了生成虚拟图像的质量, 其值越小, 生成虚拟图像属于同一类别的可能性越高.  $P(y)$  的熵值反映了生成图像的多样性, 其值越大, 生成虚拟图像的类别数越多, 训练模型的泛化能力越强. 对于泛化能力足够且生成类别均一的生成模型,  $P(y|x_G)$  的分布是尖锐的, 而  $P(y)$  是一个均匀分布. IS 值大小在一定程度上反映了生成的虚拟图像质量.

将生成的虚拟图像作为输入, 采用 ImageNet 数据集上的预训练 Inception v3 网络模型评估了图像的 IS 值, 并计算 IS 值的标准差 (STD 值). 较高的 IS 值说明 Inception 输出的平均概率分布熵值较大, 生成器生成图像的多样性较好; 较低的 STD 值说明 Inception 输出的概率分布较小, 生成图像属于同一类别的可能性较高, 图像的质量越高. DCGAN 共生成锈蚀图像 1 000 张. 随机抽样生成图像, 计算 IS 值和 STD 值, 寻找该样本用于扩增数据集的最佳数量, 结果见表 1.

由表 1 可见, 当输入图像数量为 100 ~ 400 时, 随输入图像数量的增加, IS 值不断增大, STD 值存在小幅度波动, 即图像的多样性随图像数量增加而提高, 图像的质量基本保持不变. 当输入图像数量为 500 ~ 1 000 时, 随输入图像的数量增加, IS 值存在小幅度波动, STD 值呈现下降趋势, 即图像的多样性保持在一个稳定的水平, 图像质量随图像数目增加而提高. 综合考虑生成图像的多样性和质量, <http://journal.seu.edu.cn>

表 1 DCGAN 量化分析评估值

输入图像数量	IS 值	STD 值
1 000	2.412 0	0.12
900	2.391 6	0.12
800	2.397 9	0.14
700	2.394 2	0.15
600	2.392 7	0.15
500	2.391 2	0.17
400	2.411 4	0.21
300	2.335 7	0.19
200	2.251 0	0.15
100	2.002 7	0.20

该样本用于扩增数据集的最佳数目为 800 ~ 1 000.

Mao 等<sup>[16]</sup>针对不同生成对抗网络模型在 Oxford-102 数据集上的表现进行评估, 得出不同模型的 IS 值为 2.66 ~ 3.95; 张光华等<sup>[17]</sup>针对生成的恒星和星系图像进行评估, IS 值约为 0.30. 本文选用的 DCGAN 在生成虚拟锈蚀数据集上的 IS 值为 2.41, 考虑到图像本身特征的差异性和原始数据集多样性的差距, DCGAN 在锈蚀生成上表现良好.

## 2.2 K 均值聚类算法评估分析

采用无监督的分类方法对原始数据集和生成数据集进行比较, 得到分类个数并观察结果. 如果二者分类结果一致, 说明生成虚拟图像数据集无类别损失, 可用于数据集扩增.

无监督分类包括层次聚类<sup>[18]</sup>、K 均值聚类算法、EM 算法<sup>[19]</sup>等. 考虑对图像聚类评价效果的适用性, 本文首先用 SIFT 算子提取图像特征, 并选用 K 均值聚类算法进行分类, 得到类别数. 绘制类别总误差随类别数变化的肘形图, 结果见图 4. 为避免各特征之间数值差异过大导致部分特征对模型贡献度过低的问题, 在分类前对原始数据和生成数据均进行归一化处理. 由图可见, 随着类别数的增大, 类别总误差迅速下降; 当类别数大于临界值 5 后, 类别总误差趋于稳定. 临界值 5 即为最佳的类别数.

综上所述, 当聚类数小于 5 时, 图像类别分类较少, 类别总误差较大. 当聚类数大于 5 时, 图像类别分类较多, 类别总误差较小, 但是类别之间相似度较高. 无论原始数据和生成数据归一化与否, 图的拐点/肘点都在 5 左右, 即原始数据和生成数据都不存在奇异数据, 且总体有效类别数均为 5, 二者吻合. 这说明生成数据较好地模拟了原始数据集, 生成的虚拟数据集可用于数据集扩增.

## 3 锈蚀病害分割算法的对比分析

### 3.1 基于 U-Net 网络的锈蚀区域分割

锈蚀区域分割是指划分出图像中锈蚀的准确

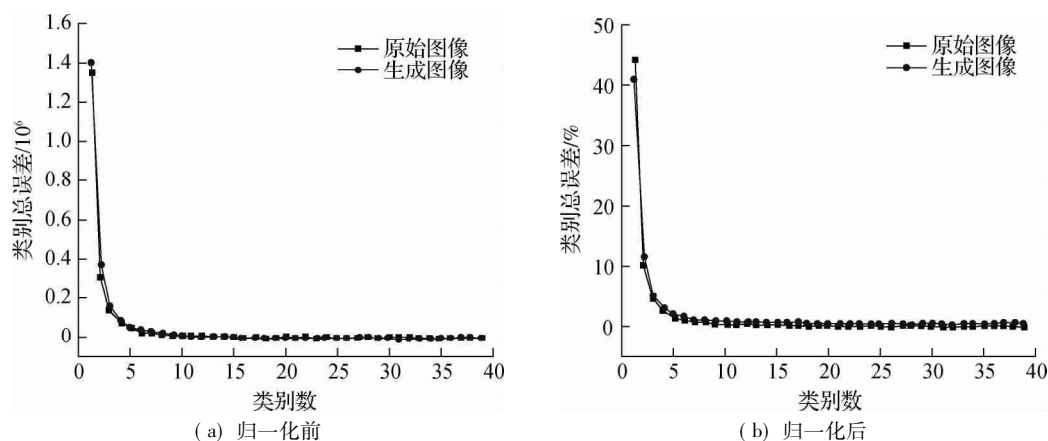


图4 图像类别总误差与类别数的关系肘形图

轮廓. 相比于另外经典的语义分割网络 FCN 和 SegNet, U-Net 网络具有训练数据量少且精度高的特点<sup>[20]</sup>. U-Net 网络由主干特征提取网络与特征融合网络组成, 其网络结构见图 5. 在主干特征提取网络中, 通过  $3 \times 3$  卷积、ReLU 激活函数及  $2 \times 2$

最大池化, 逐步对特征图进行下采样, 使得特征图尺寸逐渐减小, 通道数逐渐增加<sup>[21-22]</sup>. 在特征融合网络中, 对特征图进行  $2 \times 2$  转置卷积以及  $1 \times 1$  卷积, 逐步恢复图像细节, 使得特征图的尺寸逐渐增加, 通道数逐渐减少, 最终输出语义分割图像.

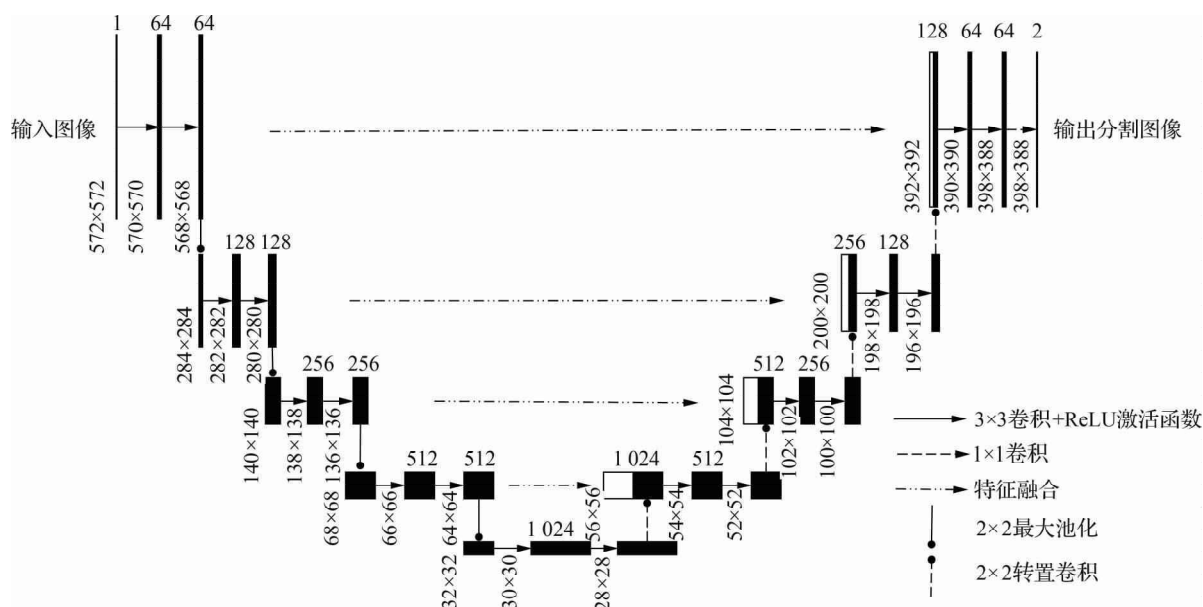


图5 U-Net 网络结构

U-Net 网络在卷积层之间加入 ReLU 激活函数<sup>[23]</sup>, 即

$$f(u) = \begin{cases} u & u \geq 0 \\ 0 & u < 0 \end{cases} \quad (7)$$

式中,  $u$  和  $f(u)$  分别为 ReLU 激活函数的输入值和输出值.

ReLU 激活函数曲线如图 6 所示. ReLU 激活函数具有单侧一致特性. 相较于 sigmoid 函数和 tanh 函数, 卷积神经网络经过 ReLU 激活函数后, 参数间的依存关系更少. 同时, ReLU 激活函数提高了网络模型的泛化性, 使计算过程简单, 且解决

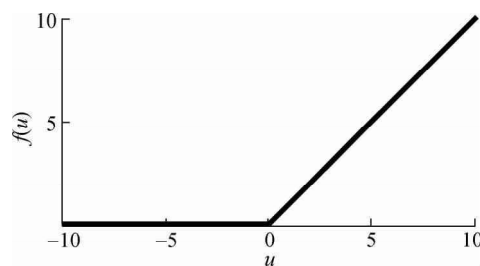


图6 ReLU 激活函数曲线

了梯度消失的问题<sup>[24]</sup>.

将 U-Net 网络和阈值分割方法相结合, 可以实现锈蚀分割. 首先, 输入彩色图像, 利用 U-Net 网络

<http://journal.seu.edu.cn>



进行训练与验证,得到灰度图(包括特征灰度图和待分割灰度图).然后,在灰度图上选取合适的灰度阈值,分割得到二值图.在超参数设置中,训练批次大小设置为 8,训练迭代轮次设置为 6 800,学习率设置为 0.01.训练过程中损失值的变化曲线见图 7.由图可知,在前 1 000 次迭代轮次中,损失值波动较大但整体收敛,之后保持稳定.最终训练的模型在验证集上运行,锈蚀分割结果见图 8.图中红色区域为锈蚀标注区域.

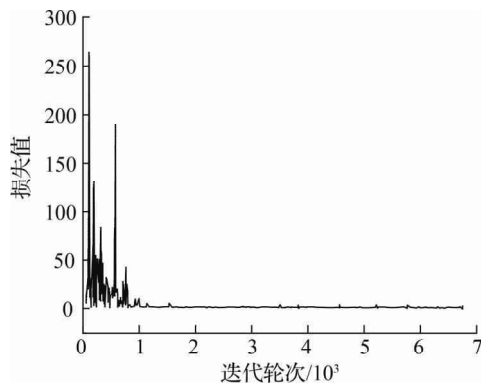


图 7 U-Net 网络训练的损失值变化曲线

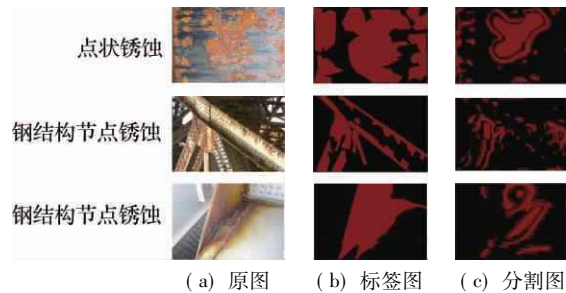


图 8 U-Net 网络模型在验证集上分割结果

由图 8 可知,对于以点状锈蚀为主的原图,U-Net 网络模型基本分割出左侧点状锈蚀,右侧点状锈蚀区域漏识别率高.对于钢结构节点锈蚀图,U-Net 网络模型能大体分割出锈蚀区域,但在轮廓和边界上分割效果一般.

### 3.2 基于 DeepLab-V3 + 网络的锈蚀区域分割

DeepLab 网络是经典的语义分割网络之一,已迭代几代网络结构,包括 V1、V2、V3 以及 V3 + 网络<sup>[25-26]</sup>.DeepLab-V3 + 网络由 1 个编码器与 1 个解码器组成(见图 9).

编码器包括 Xception 模块和 ASSP 模块.

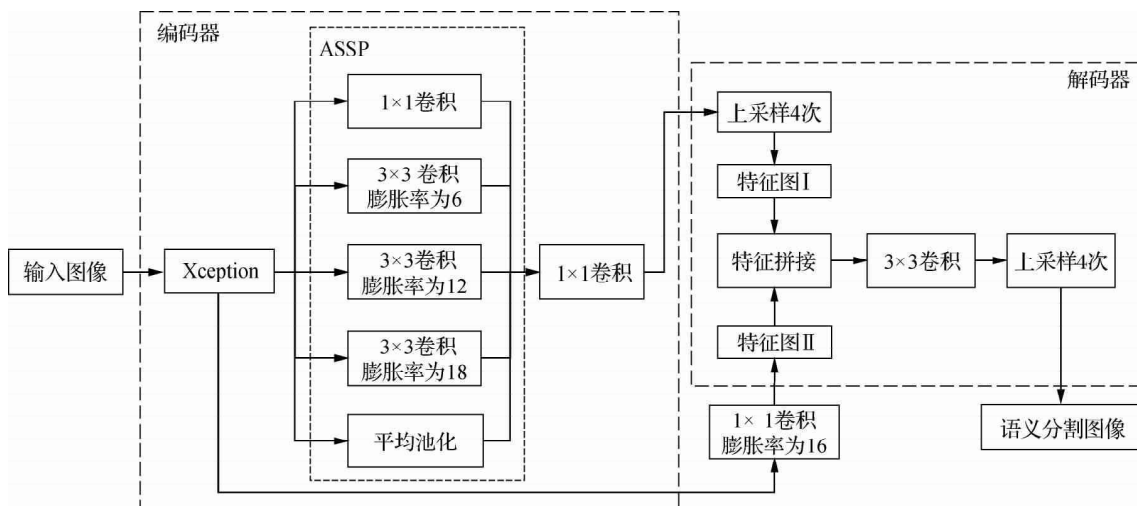


图 9 DeepLab-V3 + 网络结构

Xception 模块采用深度可分离卷积及残差网络,减少卷积参数量,加快网络收敛速度. ASSP 模块将 Xception 模块的特征图作为输入,分别进行  $1 \times 1$  卷积、3 个  $3 \times 3$  卷积(膨胀率分别为 6、12 及 18)以及平均池化.然后,将上述卷积和池化后的特征图拼接后进行  $1 \times 1$  卷积,送入解码器.

在解码器中,将 ASSP 模块输出的特征图进行 4 次上采样得到特征图 I,将 Xception 模块输出的特征图进行  $1 \times 1$  卷积(膨胀率为 16)得到特征图 2.特征图 1 和特征图 II 进行特征拼接后,进行  $3 \times 3$  卷积和 4 次上采样,得到语义分割图像.

<http://journal.seu.edu.cn>

### 3.3 实验结果及分析

本文算法所采用的编程语言为 Python3.8,网络搭建基于 PyTorch 开源框架,版本为 1.10.硬件配置如下:CPU 型号为 i7-8700k,显卡型号为 NVIDIA GTX1080ti,显存为 10 GB,操作系统为 Windows10.

数据集共包含 1 500 张图像和对应的掩码图像,图像分辨率为  $512 \times 512$  像素.数据集由原始锈蚀图像和 DCGAN 生成的虚拟锈蚀图像组成.网络采用 PyTorch 框架进行训练,将数据集随机划分为训练集和验证集,两者数量比例为 4:1.设置训练

参数,学习率为 0.007,训练迭代轮次为 150,采用如下的二分类交叉熵损失函数:

$$L_1 = - \sum_{j=1}^m (S_j \ln \hat{S}_j + (1 - S_j) \ln(1 - \hat{S}_j)) \quad (8)$$

式中,  $m$  为图像的像素点总数;  $\hat{S}_j$  表示第  $j$  个像素点预测为锈蚀点的概率值;  $S_j$  表示第  $j$  个像素点的真实标签. 若像素点为锈蚀点,则  $S_j = 1$ ; 若像素点为背景,则  $S_j = 0$ .

DeepLab-V3 + 网络训练的损失值变化曲线见图 10. 由图可见,随着迭代轮次的增加,损失值平稳逐渐下降并最终收敛.

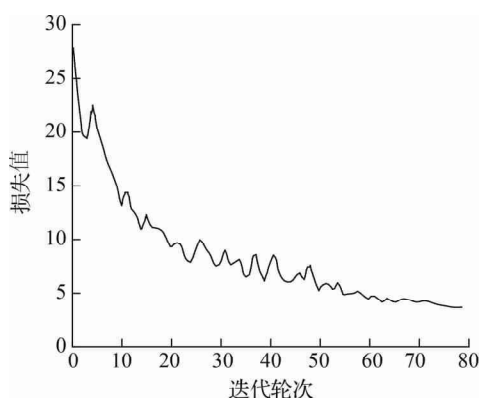


图 10 DeepLab-V3 + 网络训练的损失值变化曲线

衡量网络模型的指标包括精度 ( $P_{\text{PRE}}$ )、召回率 ( $R_{\text{REC}}$ ) 及正确率 ( $A_{\text{ACC}}$ ), 计算公式如下:

$$P_{\text{PRE}} = \frac{Q_{\text{TP}}}{Q_{\text{TP}} + Q_{\text{FP}}} \quad (9)$$

$$R_{\text{REC}} = \frac{Q_{\text{TP}}}{Q_{\text{TP}} + Q_{\text{FN}}} \quad (10)$$

$$A_{\text{ACC}} = \frac{Q_{\text{TP}} + Q_{\text{TN}}}{Q_{\text{TP}} + Q_{\text{TN}} + Q_{\text{FP}} + Q_{\text{FN}}} \quad (11)$$

式中,  $Q_{\text{TP}}$  为分类正确的正类别数;  $Q_{\text{TN}}$  为分类正确的负类别数;  $Q_{\text{FP}}$  为分类错误的正类别数;  $Q_{\text{FN}}$  为分类错误的负类别数.

$F_1$  分数是一种同时考虑模型精度和召回率的指标,其计算公式为

$$F_1 = 2 \frac{P_{\text{PRE}} R_{\text{REC}}}{P_{\text{PRE}} + R_{\text{REC}}} \quad (12)$$

对语义分割模型进行精度评估时,每个像素点的分类结果都对模型的  $F_1$  分数产生影响. 因此,  $F_1$  分数可以较为全面地评价模型的分割效果. 表 2 给出了网络模型精度评价指标.

表 2 网络模型精度评价指标

网络模型	精度	召回率	$F_1$ 分数
U-Net	0.737	0.682	0.708
DeepLab-V3 +	0.935	0.952	0.943

由表 2 可见, DeepLab-V3 + 网络模型在验证集上精度为 0.935, 召回率为 0.952,  $F_1$  分数为 0.943. 相较于 U-Net 网络模型, 精度提升了 0.198, 召回率提升了 0.270,  $F_1$  分数提高了 0.235. 由此可知, DeepLab-V3 + 网络模型具有更高的锈蚀分割精度.

将训练后的 DeepLab-V3 + 及 U-Net 网络模型对钢结构锈蚀区域进行分割, 结果见图 11. 图中红色区域为锈蚀标注区域. 由图可知, U-Net 模型因为具有较多的池化操作而丢失空间特征信息, 未能将锈蚀区域全部分割出. 对于以点状锈蚀为主的原图, DeepLab-V3 + 网络模型可有效连通点状锈蚀区域; 对于以片状锈蚀为主的原图, DeepLab-V3 + 网络模型可有效分割出片状锈蚀区域. DeepLab-V3 + 网络模型的级联空洞卷积保存多次下采样的位置信息, 因而对点状和片状锈蚀分割均具有更高的精度和算法鲁棒性.

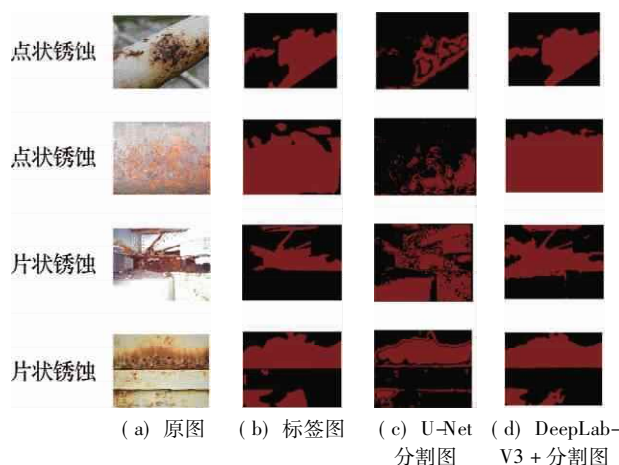


图 11 DeepLab-V3 + 及 U-Net 网络模型分割锈蚀区域图

## 4 结论

1) DCGAN 生成虚拟锈蚀数据集的 IS 值达到 2.41, 满足图像质量要求. 同时, 虚拟锈蚀数据集的分类肘形图类别数为 5, 与原数据集一致, 没有类别损失.

2) 针对锈蚀语义分割, DeepLab-V3 + 网络模型精度为 0.935, 召回率为 0.952,  $F_1$  分数为 0.943. 相较于 U-Net 网络模型, 精度提升了 0.198, 召回率提升了 0.270,  $F_1$  分数提高了 0.235, 均显著高于 U-Net 网络模型.

3) 对于片状锈蚀区域, DeepLab-V3 + 网络模型与 U-Net 网络模型均取得了较为理想的分割结果. 对于点状锈蚀区域, DeepLab-V3 + 网络模型将

<http://journal.seu.edu.cn>

点状锈蚀区域进行有效连通,分割精度显著高于 U-Net 网络模型. DeepLab-V3 + 网络可为桥梁锈蚀病害的智能语义分割提供技术支撑.

### 参考文献 (References)

- [1] 左志亮,张帆,罗赤字,等. 缓黏结预应力混凝土梁耐久性能试验研究[J]. 土木工程学报,2019,52(9): 69 - 78. DOI: 10.15951/j.tmgcxb.2019.09.005.  
Zuo Z L, Zhang F, Luo C Y, et al. Research on durability of retard-bonded prestressed concrete beams[J]. *China Civil Engineering Journal*, 2019, **52**(9): 69 - 78. DOI: 10.15951/j.tmgcxb.2019.09.005. (in Chinese)
- [2] 朱劲松,李欢. 基于深度学习的钢桥病害分割与量化[J]. 东南大学学报(自然科学版),2022,52(3): 516 - 522. DOI: 10.3969/j.issn.1001-0505.2022.03.012.  
Zhu J S, Li H. Segmentation and quantification of steel bridge defects based on deep learning[J]. *Journal of Southeast University (Natural Science Edition)*, 2022, **52**(3): 516 - 522. DOI: 10.3969/j.issn.1001-0505.2022.03.012. (In Chinese)
- [3] Petricca L, Moss T, Figueroa G, et al. Corrosion detection using AI: A comparison of standard computer vision techniques and deep learning model[C]//*Proceedings of the Sixth International Conference on Computer Science, Engineering and Information Technology*. Vienna, Austria, 2016, **91**: 2 - 16. DOI: 10.5121/csit.2016.60608.
- [4] Deegan J A, Mohammad R J. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection[J]. *Structural Health Monitoring*, 2018, **17**(5): 1110 - 1128. DOI: 10.1177/1475921717737051.
- [5] 王达磊,彭博,潘玥,等. 基于深度神经网络的锈蚀图像分割与定量分析[J]. 华南理工大学学报(自然科学版),2018,46(12): 121 - 127,146. DOI: 10.3969/j.issn.1000-565X.2018.12.015.  
Wang D L, Peng B, Pan Y, et al. Segmentation and quantitative analysis of corrosion images based on deep neural networks[J]. *Journal of South China University of Technology (Natural Science Edition)*, 2018, **46**(12): 121 - 127, 146. DOI: 10.3969/j.issn.1000-565X.2018.12.015. (in Chinese)
- [6] Chen Y J, Qi L, Sun H Y, et al. Metal surface corrosion grade estimation from single image[C]//*Ninth International Conference on Graphic and Image Processing*. Qingdao, China, 2018: 1404 - 1409. DOI: 10.1117/12.2302776.
- [7] Bastian B T, Jaspreeth N, Ranjith S K, et al. Visual inspection and characterization of external corrosion in pipelines using deep neural network[J]. *NDT and E International*, 2019, **107**: 102134. DOI: 10.1016/j.ndteint.2019.102134.
- [8] Qian C. *Evaluation of deep learning-based semantic segmentation approaches for autonomous corrosion detection on metallic surfaces* [D]. West Lafayette, IN, USA, Purdue University, 2019.
- [9] Xu J, Gui C, Han Q. Recognition of rust grade and rust ratio of steel structures based on ensembled convolutional neural network[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2020, **35**(10): 1160 - 1174. DOI: 10.1111/mice.12563.
- [10] 杨耀坤,肖卫国,王力. 基于生成式对抗网络的小样本图像生成[J]. 计算机仿真,2021,38(12): 445 - 448,453. DOI: 10.3969/j.issn.1006-9348.2021.12.091.  
Yang Y K, Xiao W G, Wang L. Research on image generation method using small samples based on generative adversarial networks[J]. *Computer Simulation*, 2021, **38**(12): 445 - 448, 453. DOI: 10.3969/j.issn.1006-9348.2021.12.091. (in Chinese)
- [11] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. *Advances in Neural Information Processing Systems*, 2014, **27**: 2672 - 2680. DOI: 10.48550/arXiv.1406.2661.
- [12] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//*2016 International Conference on Learning Representations*. San Juan, Puerto Rico, 2016: 1 - 15. DOI: 10.48550/arXiv.1511.06434.
- [13] Salimans T, Goodfellow I J, Zaremba W, et al. Improved techniques for training GANs[C]//*Proceedings of the 30th International Conference on Neural Information Processing Systems*. Red Hook, NY, USA, 2016: 2234 - 2242. DOI: 10.48550/arXiv.1606.03498.
- [14] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium[C]//*Proceedings of the 31th International Conference on Neural Information Processing Systems*. Red Hook, NY, USA, 2017: 6629 - 6640. DOI: 10.48550/arXiv.1706.08500.
- [15] MacQueen J. Some methods for classification and analysis of multivariate observations[C]//*Proc 5th Berkeley Symposium on Math, Stat, and Prob*. Berkeley, CA, USA, 1965: 281 - 297.
- [16] Mao F L, Ma B P, Chang H, et al. Learning efficient text-to-image synthesis via interstage cross-sample similarity distillation[J]. *Science China (Information Sciences)*, 2021, **64**(2): 19 - 30. DOI: 10.1007/s11432-020-2900-x.
- [17] 张光华,王福豹,段渭军. 一种针对 GAN 生成的天文

http://journal.seu.edu.cn



- 图像评价方法研究[J]. 计算机仿真, 2020, 37(1): 464-468. DOI: 10.3969/j.issn.1006-9348.2020.01.097.
- Zhang G H, Wang F B, Duan W J. A method for evaluating astronomical images generated by GAN[J]. *Computer Simulation*, 2020, 37(1): 464-468. DOI: 10.3969/j.issn.1006-9348.2020.01.097. (in Chinese)
- [18] Ward Jr J H. Hierarchical grouping to optimize an objective function[J]. *Journal of the American Statistical Association*, 1963, 58(301): 236-244. DOI: 10.1080/01621459.1963.10500845.
- [19] Dempster A P, Laird N M, Rubin D B. Maximum Likelihood from incomplete data via the EM algorithm[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, 39(1): 1-22. DOI: 10.1111/j.2517-6161.1977.tb01600.x.
- [20] Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(5): 749-753. DOI: 10.1109/LGRS.2018.2802944.
- [21] 陈法法, 成孟腾, 杨蕴鹏, 等. 融合双注意力机制和U-Net网络的锈蚀图像分割[J]. 西安交通大学学报, 2021, 55(12): 119-128. DOI: 10.7652/xj-tuxb202112014.
- Chen F F, Cheng M T, Yang Y P, et al. A segmentation method based on dual attention mechanism and U-Net for corrosion images[J]. *Journal of Xi'an Jiaotong University*, 2021, 55(12): 119-128. DOI: 10.7652/xjtuxb202112014. (in Chinese)
- [22] 周莉莉, 姜枫. 图像分割方法综述研究[J]. 计算机应用研究, 2017, 34(7): 1921-1928. DOI: 10.3969/j.issn.1001-3695.2017.07.001.
- Zhou L L, Jiang F. Survey on image segmentation methods[J]. *Application Research of Computers*, 2017, 34(7): 1921-1928. DOI: 10.3969/j.issn.1001-3695.2017.07.001. (in Chinese)
- [23] Li X J, Qian W H, Xu D, et al. Image segmentation based on improved U-Net[J]. *Journal of Physics: Conference Series*, 2021, 1815(1): 012018. DOI: 10.1088/1742-6596/1815/1/012018.
- [24] 周正东, 李剑波, 辛润超, 等. 基于带孔U-Net神经网络的肺癌危及器官并行分割方法[J]. 东南大学学报(自然科学版), 2019, 49(2): 231-236. DOI: 10.3969/j.issn.1001-0505.2019.02.005.
- Zhou Z D, Li J B, Xin R C, et al. Parallel segmentation method for organs at risk in lung cancer based on dilated U-Net neural network[J]. *Journal of Southeast University (Natural Science Edition)*, 2019, 49(2): 231-236. DOI: 10.3969/j.issn.1001-0505.2019.02.005. (in Chinese)
- [25] Papandreou G, Chen L C, Murphy K P, et al. Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation[C]// *Proceedings of the IEEE International Conference on Computer Vision*. Santiago, Chile, 2016: 1742-1750. DOI: 10.1109/ICCV.2015.203.
- [26] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834-848. DOI: 10.1109/TPAMI.2017.2699184.