

基于生成对抗网络的飞机舱门图像涂装增广方法

李 旭¹, 黄川峻²

(1. 北京博维航空设施管理有限公司, 北京 100621; 2. 电子科技大学 自动化工程学院, 成都 611731)

摘要: 舱门检测网络是自动驾驶登机桥检测舱门的重要实现途径; 而在舱门图像带有复杂涂装的场景下, 舱门检测网络对舱门图像的检测准确率、召回率和 mAP 大幅降低; 针对这一问题, 提出了基于生成对抗网络的舱门图像涂装增广方法, 用以提高舱门检测网络训练集中带复杂涂装的图像样本的数量, 从而提高舱门检测网络在复杂涂装场景下的检测准确率、召回率和 mAP ; 通过将待增广图像的边缘图像和涂装图像的边缘图像进行边缘融合, 再经过已训练的生成对抗网络填色, 实现舱门图像数据集的涂装增广; 实验证明, 与传统的基础增广方法相比, 基于生成对抗网络的舱门图像涂装增广方法对 YOLOv5s 的检测准确率、召回率和 mAP 提升更高。

关键词: 舱门图像; 图像增广; 生成对抗网络; 涂装增广; 目标检测

Painting Augmentation Method for Aircraft Door Images Based on Generative Adversarial Network

LI Xu¹, HUANG Chuanjun²

(1. Beijing Bowei Aviation Facilities Management Co., Ltd., Beijing 100621, China; 2. School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: The aircraft door detection network is an important way for autonomous boarding bridges to detect aircraft doors. However, in the scenario where the aircraft door image has complex painting, the door detection network greatly reduces the detection accuracy, recall rate and mAP of door images. To address this problem, a door image painting augmentation method based on generative adversarial network is proposed to increase the number of image samples with complex painting in the door detection network training set, thereby improving the detection accuracy, recall rate and mAP of the network in complex painting scenes. By fusing the edge image from the augmented edge image and the painted edge image, and then coloring through the trained generative adversarial network, the painting augmentation of the door image dataset is achieved. Experimental results show that compared with traditional basic augmentation methods, the door image painting augmentation method based on the generative adversarial network has a higher improvement in the detection accuracy, recall rate and mAP of YOLOv5s.

Keywords: aircraft door image; image augmentation; generative adversarial network; painting augmentation; object detection

0 引言

随着“智慧机场”的提出和兴起, 登机桥自动对接已成为国内客运机场未来发展的一大趋势。其中, 登机桥自动对接是依赖登机桥自动驾驶系统实现的。舱门检测网络是自动驾驶登机桥检测舱门的重要实现途径。而在舱门图像带有复杂涂装的场景下, 舱门检测网络对舱

门图像的检测准确率、召回率和 mAP 大幅降低。如图 1 中展示了一副涂装较为复杂的舱门图像。相较于普通的没有复杂涂装的舱门图像, 舱门检测网络对于这类具有复杂涂装的舱门图像的检测准确率、召回率和 mAP 较低。舱门检测网络通常基于目标检测神经网络构建。而用于目标检测的神经网络的性能指标高低与训练数据集中的目标种类的数据集数量和多样性有着密切的联

收稿日期: 2024-06-03; 修回日期: 2024-08-07。

基金项目: 国家自然科学基金(62073196)。

作者简介: 李 旭(1985-), 男, 大学本科, 工程师。

通讯作者: 黄川峻(1998-), 男, 硕士。

引用格式: 李 旭, 黄川峻. 基于生成对抗网络的飞机舱门图像涂装增广方法[J]. 计算机测量与控制, 2025, 33(7): 188-194, 202.

系。而在实际工程实践中,考虑到机场管制、带复杂涂装的舱门图像较为稀有等因素,复杂涂装舱门图像的采集存有较大的困难,这也导致了往往舱门检测网络训练数据集的复杂涂装舱门图像数量的不足,进而使得舱门检测网络对复杂涂装舱门图像的检测能力的不足。



图1 复杂涂装舱门图像

针对舱门检测网络训练数据集中复杂涂装舱门图像的数量和多样性不足导致的舱门检测网络对复杂涂装舱门图像的检测准确率、召回率和 mAP 较低的问题,拟对检测网络训练数据集的复杂舱门图像样本数量进行增广,用以提升舱门检测网络在检测复杂涂装舱门图像场景下的检测准确率、召回率和 mAP 。

按照是否使用神经网络,现有的图像增广方法可以划分为传统的基础图像增广方法和基于神经网络的图像增广方法。

传统的图像增广方法是一类通过对增广过程进行人工建模,从而实现对图像数据集进行扩充的增广方法。根据建模方式的不同,传统的图像增广方法又包括几何变换增广、噪声插入增广^[1]、模糊增广、色彩空间变换增广、随机擦除增广^[2]、裁剪组合增广^[3]、混合增广^[4]等。传统的图像增广凭借着方法简单、适用广泛的特点,被广泛应用与各种训练任务的数据集增广中。但是,传统的图像增广方法无法完成数据集样本在某些特定复杂场景(例如:使舱门带有涂装)的增广任务,对所增广的网络的提升效果有限。随着计算机算力的提高和大数据集的应用,神经网络凭借其自学习性、自组织性、自适应性得到大量学者的关注。一些学者将神经网络引入图像增广领域,提出了众多基于神经网络的图像增广方法。

根据增广操作对象的不同,基于神经网络的图像增广方法包括特征空间增广和图像空间增广。大部分的图像增广方法的操作对象均是待增广图像的图像空间,部分学者将图像的特征空间作为图像增广方法的操作对象。文献[5]发现,通过对图像的中间层单独进行增广操作,网络在 CIFAR-100 的性能从 66% 提升到 73%。2017 年,文献[6]提出了一种简单的、与图像域无关的、基于图像特征空间的图像增广方法。尽管特征空间增广为图像增广提供了一种全新的方向,但是以特征向量为对象的增广无法按照指定的要求增广图像,

这种增广方法无法胜任图像翻译问题。

风格迁移增广是基于神经网络的图像空间增广的一个重要组成部分。风格迁移是计算机视觉的重要任务,可以将输入图像从一个样式转化成另一种样式。近年来,风格迁移也被用于图像增广领域。

2015 年,文献[7]在卷积神经网络(CNN, Convolutional Neural Network)^[8]基础上提出了一种基于优化的风格迁移网络。该网络生成的风格迁移图像在视觉上取得了令人信服的效果,一定程度上解决了传统的风格迁移算法稳定性不佳、效果不佳的问题。但所采用的网络结构和基于优化的方法也使得该风格迁移网络生成图像速度较慢,并且每个训练网络完成的该网络只能完成单幅内容图像单种风格的迁移,无法实现多幅内容图像多风格的迁移,更加无法用于图像增广等对速度和泛用性有一定要求的任务。为解决这一问题,文献[9]、文献[10]、文献[11]相继提出了各自的实时风格迁移网络,在图像生成速度上比文献[7]快了 3 个数量级。紧接着,文献[12]通过对两种风格进行加权求和的方式证实了风格嵌入的线性可叠加性。2019 年,文献[13]在文献[7]的基础上提出了一种基于随机风格迁移的新的数据增广方法。实验表明,文献[13]的风格迁移增广方法可以在较短时间对图像数据集进行大规模的扩充,并且该方法还是域无关的,可以用于所有种类数据集的增广。

基于神经网络的图像增广方法凭借其强大的自建模能力正在成为图像增广领域未来的趋势,基于神经网络的图像增广所依赖的增广网络的训练本身也需要大量的训练样本的训练,这与图像增广的初衷之一即解决样本数量不足存在一定的矛盾性。此外,基于神经网络的图像增广方法目前尚还没有针对在无复杂涂装的舱门图像增加涂装的增广方法。如何实现舱门图像的涂装增广存在不小的难度。

1 基于生成对抗网络的舱门图像涂装增广

本文对舱门图像的涂装分布情况划分如图 2 所示。按照舱门区域是否存在涂装以及涂装的复杂程度,将舱门图像划分为无涂装舱门图像、简单涂装舱门图像和复杂涂装舱门图像。无涂装舱门图像指舱门区域几乎没有涂装覆盖的舱门图像。简单涂装舱门图像指舱门区域带有少量涂装的舱门图像。复杂涂装舱门图像指舱门区域带有大量涂装的舱门图像,这类涂装往往出于广告宣传等目的单独设计,涂装大多具有使用时间较短、种类丰富等特点,在构图设计上一般较为复杂。对于这类舱门图像,由于样本数量少、对舱门区域的干扰大,舱门检测网络对这类图像误检、漏检的概率远远高于另外两类舱门图像。

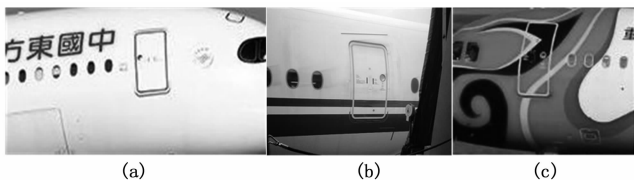


图 2 舱门区域涂装覆盖情况

由于传统的图像增广方法无法完成对舱门检测网络的训练数据集复杂涂装舱门样本数量的有效扩充，这些方法对舱门检测网络在复杂涂装舱门图像的准确率、召回率和 mAP 检测指标的提升也比较有限。因此，如何扩充检测网络训练集中的复杂涂装图像样本数量是本文的一个重点。而舱门检测网络的训练数据集中缺少带复杂涂装的舱门图像，如何在缺少复杂涂装舱门图像样本的数据集中生成复杂涂装舱门图像样本是本文的一大难点。

如果忽略舱门图像的背景部分，大部分舱门图像可以抽象为两个部分，分别是舱门图像的边缘图像和舱门图像边缘之间的色块部分。因此，舱门图像可以看作是边缘图像和边缘之间的色块组成的像素集合。基于这个想法，受草图生成图像的风格迁移任务的启发，舱门图像或许可以由边缘图像生成。

因此，为了在缺少复杂涂装的舱门图像数据集中增广得到带复杂涂装的舱门图像，如图 3 所示，本文提出了一种基于舱门—边缘—舱门过程的舱门图像涂装增广过程。首先，通过边缘提取算法提取待增广舱门图像的边缘图像，并通过同样的方式从涂装图像中提取得到涂装边缘图像。然后，通过边缘融合将舱门图像的边缘图像和涂装边缘图像进行融合，得到一张边缘融合图像。最后，将得到的边缘融合图像送入已训练的增广网络的生成器中，从而得到最终的涂装增广图像。

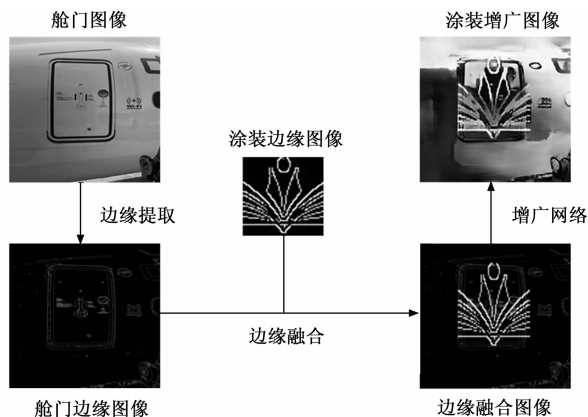


图 3 舱门图像涂装增广过程

增广网络的主要作用是将边缘融合图像转化为涂装增广图像，因此，增广网络的输入和输出均是图像。并且增广网络的输出图像不能够像原始的生成对抗网络一

样随机产生图像，而是要根据输入图像去产生输出图像。本文选择具备能够训练成对图像的网络作为增高光网络，常见的满足条件的主流神经网络有 Pix2pix^[14] 网络、CycleGAN^[15] 网络和 SketchGAN^[16] 网络等。

本文设计的舱门涂装增广方法如图 4 所示。舱门涂装增广方法分为两个过程，分别是训练过程和增广过程。训练过程的主要目标是训练得到一个可以由边缘图像生成彩色舱门图像的增广网络，用于扩充待增广数据集中的带复杂涂装的舱门图像。增广过程的主要目标是通过已训练的增广网络的生成器将边缘融合图像转换为涂装增广图像，实现对待增广数据集中的带复杂涂装的舱门图像数量的扩充。

如图 4 (a) 所示，训练过程中的增广网络由两部分组成，分别为生成器和判别器。每个样本的训练过程中，舱门图像通过边缘提取得到对应的舱门边缘图像。随后，将舱门边缘图像送入增广网络的生成器中得到生成图像。紧接着，将舱门图像和生成图像一并送入判别器分别得到舱门图像的预测值和生成图像的预测值，用于计算得到生成器和判别器权重更新迭代所需要的损失值。舱门涂装增广方法的增广过程如图 4 (b) 所示，每次增广过程中，待增广图像通过边缘提取得到待增广图像的舱门边缘图像。随后，将原始图像的边缘图像和涂装边缘图像进行边缘融合得到边缘融合图像。最后，将边缘融合图像作为输入送入已训练的增广网络生成器中，从而产生涂装增广图像，进而完成对待增广数据集中的带复杂涂装的舱门图像数量的扩充。

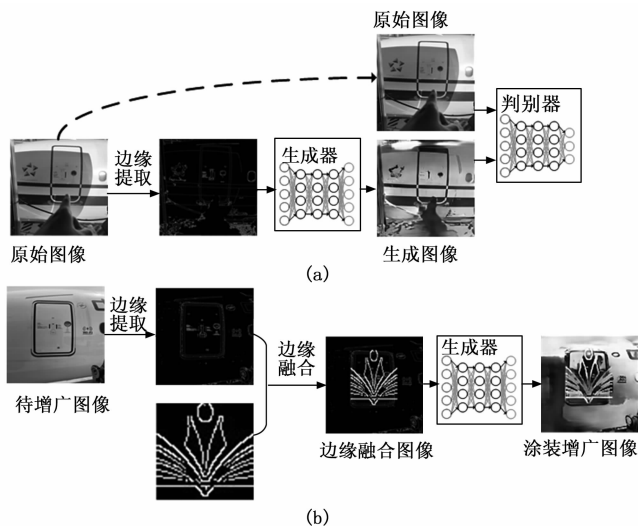


图 4 舱门涂装增广方法

本章节设计的舱门图像涂装增广方法步骤如下：

- 1) 建立边缘/舱门配对训练集和涂装图像数据集，用于增广网络的训练。边缘/舱门配对数据集是通过将涂装舱门图像和它们的边缘图像配对建立的。
- 2) 使用边缘/舱门配对训练集训练增广网络，用于

图像增广。

3) 提取待增广舱门图像和涂装图像的边缘图像, 用于边缘融合。

4) 融合待增广舱门图像的边缘图像和涂装图像的边缘图像, 得到边缘融合图像, 用于作为增广网络增广时的输入图像。

5) 将融合边缘图像输入到已训练的增广网络, 得到涂装增广图像。

2 生成图像实验与分析

2.1 检测指标

FIDScore: FIDScore^[17] (Frechet Inception Distance Score) 是一种常用的生成对抗网络评价指标, 用于衡量生成图像集和真实图像集的接近程度。对于两个图像集, FIDScore 越小意味着两个图像数据集越接近。FIDScore 等于 0 也就意味着两个图像数据集完全相同。因此, 生成图像集和真实图像集的 FIDScore 越接近于 0, 则意味着生成的图像越接近于真实图像。

PSNR: 峰值信噪比 PSNR (Peak Signal to Noise Ratio) 是衡量图像质量的指标之一。PSNR 衡量的最大像素值和均方误差之间的比值, PSNR 越大, 则图像质量越好。

SSIM: 结构相似性 SSIM (Structural Similarity) 是一种用于评估两幅图像相似度的指标, 常用于衡量图像失真前与失真后的相似性, 也用于衡量模型生成图像的真实性, 如图像去雨、图像去雾等。SSIM 越高, 说明两幅图像的结构相似性越强。

2.2 实验设计

舱门涂装增广的增广网络的训练过程如图 5 所示。增广网络选用增广网络训练集进行训练。增广网络分别由生成器和判别器组成。用于训练的增广网络训练集的每组样本均由真实图像和由真实图像对应的边缘图像组成。每组样本的训练过程中, 边缘图像作为输入送入生成器产生生成图像, 并和真实图像一并送入判别器产生损失值。通过损失值反过来更新生成器与判别器的参数。

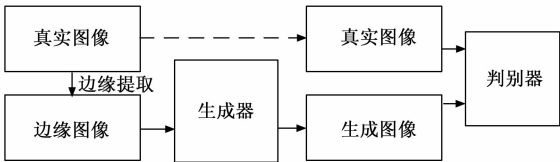


图 5 舱门涂装增广网络训练示意图

实验抽取 1 498 张边缘融合图像对测试网络原始训练集进行增广, 进而送入已训练的增广网络生成 1 498 张增广图像。增广网络训练 200 轮, 前 100 轮保持 0.000 2 的学习率, 后 100 轮学习率由 0.000 2 线性减少至 0。

本文还设计了舱门图像涂装增广方法的对比实验。本次实验分别使用了 Pix2pix、CycleGAN 和 SketchGAN 三种生成对抗网络作为舱门涂装增广的增广网络, 用于产生生成图像。实验通过对增广网络测试集的边缘图像和涂装边缘数据集样本进行边缘融合操作, 得到增广网络测试集的边缘融合图像。通过将边缘融合图像作为已训练增广网络的输入图像, 生成输出的舱门涂装图像。随后便可以得到不同增广网络的涂装增广图像数据集与原始的检测网络训练集之间的 FIDScore、SSIM, 并可以计算生成图像的平均 PSNR。通过不同增广网络的涂装增广数据集的 FIDScore、SSIM 和 PSNR 三项生成指标评估不同增广网络的图像生成质量。

2.3 实验数据与实验环境

本次实验所用到的数据集如表 1 所示。为训练 Pix2pix 网络由边缘图像重建舱门图像的能力, 共搜集了 4 746 张带涂装舱门图像, 并通过边缘提取得到 4 746 张边缘图像, 共同组成增广网络的成对训练数据集, 用于增广网络的训练。增广网络的部分训练样本如图 6 所示, 增广网络的训练样本均是带涂装的舱门图像, 不包含任何的无涂装的舱门图像。此外, 实验收集了 102 张涂装图像, 并通过边缘提取得到 102 张涂装边缘图像, 用于组成涂装边缘数据集。涂装边缘数据集的部分样本如图 7 所示, 这些样本均是由彩色的涂装图像进行边缘提取得到。此外本次实验还建立了增广网络测试集, 其中, 增广网络测试集共包括 1 498 张不带有复杂涂装的舱门图像, 用以作为本次实验的待增广图像。增广网络测试集的部分样本如图 8 所示, 增广网络测试集样本均是不含复杂涂装的舱门图像。

表 1 数据集信息

数据集	样本数量	是否带标签
增广网络训练集	4 746 对	否
涂装边缘数据集	102	否
增广网络测试集	1 498	是

实验环境如表 2 所示。实验环境是云端实验平台, 主要用于训练和测试增广网络。实验平台的配置是基于 Ubuntu20.04 的操作系统, GPU 是 24 GB 显存的 RTX4090, CPU 是 12 核的 Xeon (R) Platinum 8352V。实验所使用的深度学习框架是 pytorch1.8。

表 2 实验环境

参数名称	配置
GPU	RTX4090/24GB
CPU	12 核, Xeon(R) Platinum 8352V
运行系统	Ubuntu20.04

2.4 实验结果与分析

各个增广网络的生成图像生成评估指标如表 3 所



图 6 增广网络训练集



图 7 涂装边缘数据集



图 8 增广网络测试集

示。表 3 展示了不同增广网络的生成图像数据集的各项生成指标。其中，FIDScore 表示生成图像数据集与原始的舱门图像数据集之间的 FIDScore，PSNR 表示生成图像数据集的信噪比，SSIM 表示生成图像数据集与原始的舱门图像数据集的结构相似性指标 SSIM。

表 3 不同网络涂装增广数据集的生成指标比较

生成指标	Pix2pix	CycleGAN	SketchGAN
FIDScore	168.57	176.59	178.94
PSNR	16.458 6	12.354 7	11.981 7
SSIM	0.712 8	0.699 2	0.672 5

表中分别展示了不同网络作为增广网络后的增广图像。其中，Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.4586 和 0.7128。与 CycleGAN 和 SketchGAN 的指标相比，Pix2pix 的增广数据集的 FIDScore 更小，PSNR 和 SSIM 更大。FID-

Score 越小表示生成图像与原始图像特征分布越接近，PSNR 越大表示生成图像质量越高，SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。表中分别展示了不同网络作为增广网络后的增广图像。其中，Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.458 6 和 0.712 8。与 CycleGAN 和 SketchGAN 的指标相比，Pix2pix 的增广数据集的 FIDScore 更小，PSNR 和 SSIM 更大。FIDScore 越小表示生成图像与原始图像特征分布越接近，PSNR 越大表示生成图像质量越高，SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。展示了原始图像、原始图像的边缘融合图像以及对应的涂装增广图像。其中，表中分别展示了不同网络作为增广网络后的增广图像。其中，Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.458 6 和 0.712 8。与 CycleGAN 和 SketchGAN 的指标相比，Pix2pix 的增广数据集的 FIDScore 更小，PSNR 和 SSIM 更大。FIDScore 越小表示生成图像与原始图像特征分布越接近，PSNR 越大表示生成图像质量越高，SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。图 9 (a) 表示待增广数据集的原始图像样本，表中分别展示了不同网络作为增广网络后的增广图像。其中，Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.458 6 和 0.712 8。与 CycleGAN 和 SketchGAN 的指标相比，Pix2pix 的增广数据集的 FIDScore 更小，PSNR 和 SSIM 更大。FIDScore 越小表示生成图像与原始图像特征分布越接近，PSNR 越大表示生成图像质量越高，SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。图 9 (b) 表示经过边缘提取以及与涂装边缘图像融合得到的边缘融合图像，表中分别展示了不同网络作为增广网络后的增广图像。其中，Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.458 6 和 0.712 8。与 CycleGAN 和 SketchGAN 的指标相比，Pix2pix 的增广数据集的 FIDScore 更小，PSNR 和 SSIM 更大。FIDScore 越小表示生成图像与原始图像特征分布越接近，PSNR 越大表示生成图像质量越高，SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。图 9 (c) ~ (e) 表示将

边缘融合图像送入增广网络后得到的增广图像, 从左到右依次表示经过 Pix2pix 网络、CycleGAN 和 SketchGAN 生成的增广图像。由表中分别展示了不同网络作为增广网络后的增广图像。其中, Pix2pix 网络的增广图像数据集的 FIDScore、PSNR 和 SSIM 分别为 168.57、16.458 6 和 0.712 8。与 CycleGAN 和 SketchGAN 的指标相比, Pix2pix 的增广数据集的 FIDScore 更小, PSNR 和 SSIM 更大。FIDScore 越小表示生成图像与原始图像特征分布越接近, PSNR 越大表示生成图像质量越高, SSIM 越高表示生成图像与原始图像的结构相似性越高。因此可以说明 Pix2pix 网络的增广图像数据集的生成指标优于 CycleGAN 和 SketchGAN。由图 9 (c) ~ (e) 可知, 本文提出的舱门涂装增广过程成功由舱门图像的边缘图像还原得到了舱门涂装图像, 证实了本文提出的舱门涂装增广过程的可行性。

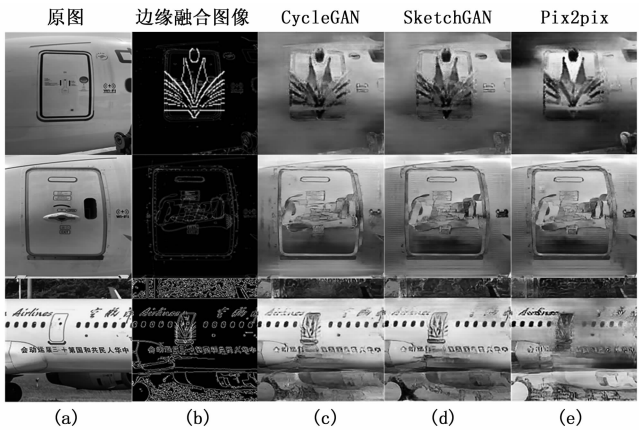


图 9 原始图像、边缘图像与增广图像展示

综上所述, 基于舱门—边缘—舱门过程的舱门涂装增广方法是可行的。通过将舱门图像转换为边缘图像, 在与涂装图像的边缘图像进行边缘融合得到边缘融合图像, 随后经过增广网络增广, 可以获得涂装增广的舱门图像。

3 检测实验与分析

3.1 检测指标

准确率和召回率。目标检测任务可以以交并比 (IoU) 衡量预测和实际的重合程度。准确率越高, 说明目标检测的误检率越低。召回率越低, 说明目标检测的漏检率越低。

mAP 是一种常用的目标检测指标, 用于衡量目标检测的质量。*mAP* 解决了准确率无法衡量漏检程度和召回率无法衡量误检程度的问题, 是一项衡量目标检测质量的综合指标。

3.2 实验设计

本次实验使用测试网络原始训练集作为 base, 并分别使用 Pix2pix、CycleGAN 和 SketchGAN 作为增广网络

对测试网络原始训练集 (待增广数据集) 进行 1:1 增广, 分别得到四组增广数据集。实验选择 YOLOv5s^[18] 和 M2Det^[19] 两种检测网络作为实验的测试网络, 用于待增广数据集和增广数据集的训练, 并在训练结束后用于测试网络测试集的检测, 从而得到在测试集上的准确率、召回率以及 *mAP* 三项检测指标。最后, 对获得的检测指标结果进行比较分析, 用于得出最后的实验结论。

3.3 实验数据与实验环境

本次实验的实验数据集信息如表 4 所示。测试网络测试集共包括 252 张复杂涂装舱门图像, 用以得到测试网络的准确率、召回率与 *mAP* 三项检测指标。测试网络原始训练集包括 1 498 张不含复杂涂装的舱门图像。涂装增广数据集由 1 498 张测试网络原始训练集经过复杂涂装增广网络 1:1 增广得到。

表 4 数据集信息

数据集	样本数量
测试网络训练集(待增广数据集)	1 498
涂装增广数据集	1 498
测试网络测试集	252

实验在基于 Ubuntu20.04 的云端实验平台进行, GPU 是 24 GB 显存的 RTX4090, CPU 是 12 核 Xeon (R) Platinum 8352V。实验所使用的深度学习架构是 pytorch1.8。

3.4 实验结果与分析

经过 YOLOv5s 和 M2Det 检测网络的测试, 未经过增广和经过不同增广网络增广得到的检测网络训练集的测试指标如表 5 所示。

表 5 不同测试网络测试不同网络增广的训练集的检测指标比较

数据集	YOLOv5s			M2Det		
	准确率 /%	召回率 /%	<i>mAP</i>	准确率 /%	召回率 /%	<i>mAP</i>
Base	79.20	81.60	0.805	77.50	79.00	0.786
Base+Pix2pix	80.10	82.70	0.813	79.00	80.20	0.818
Base+CycleGAN	80.00	82.50	0.810	78.40	80.20	0.795
Base+SketchGAN	80.40	83.10	0.819	78.80	80.70	0.799

表中分别展示了使用 YOLOv5s 和 M2Det 作为测试网络测试待增广数据集与不同增广网络增广数据集的检测指标结果。对于测试网络, 检测准确率、召回率、*mAP* 越高, 说明测试网络的检测能力越强, 出现误检、漏检的概率一般也会更低。而对于增广方法来说, 增广数据集训练的测试网络检测能力越强, 也就说明增广方法的对测试网络的提升越高。如表所示, 未经过增广的原始数据集训练的 YOLOv5s 网络的检测准确率、召回率和 *mAP* 分别是 79.20%、81.60% 和 0.805, 未经过增广的原始数据集训练的 M2Det 网络的检测准确率、召回率和 *mAP* 分别是 77.50%、79.00 和 0.786, 两种检

测网络的检测准确率、召回率和 *mAP* 三项指标均处于一个较低的水平。选择 3 种生成对抗网络 Pix2pix、CycleGAN 和 SkecthGAN 作为增广网络对待增广数据集进行涂装增广后,用以训练 YOLOv5s 和 M2det 检测网络。其中,对于 YOLOv5s,增广后的检测准确率提升了 0.8% 到 1.2%,召回率提升了 1.1% 到 1.5%,*mAP* 提升了 0.5% 到 1.4%。而对于 M2det 网络,准确率提升了 0.9% 到 1.5%,召回率提升了 1.2% 到 1.7%,*mAP* 提升了 0.9% 到 3.2%。

由此可以得出结论,在舱门检测场景下,舱门涂装增广能够提升舱门检测网络对复杂涂装舱门图像的舱门检测准确率、召回率和 *mAP*。

此外本文还设计了舱门涂装增广方法与基础增广方法的对比实验。实验分别用 Pix2pix 与 5 种基础增广方法对测试网络原始训练集 (base) 进行 1:1 增广。随后分别用 YOLOv5s 检测网络作为测试网络涂装增广数据集、5 种基础增广方法增广数据集与测试网络原始训练集,并在训练结束后用测试网络测试集检测已训练的 YOLOv5s 网络的准确率、召回率和 *mAP*。

实验共选择水平翻转、垂直翻转、模糊、随机裁剪、旋转 5 种基础增广方法,和舱门涂装增广方法一并用于 YOLOv5s 网络的训练中。YOLOv5s 测试涂装增广方法和基础增广方法的检测指标比较如表 6 所示。

表 6 涂装增广方法和基础增广方法的检测指标比较

数据集	YOLOv5s		
	准确率/%	召回率/%	<i>mAP</i>
Base	79.2	81.6	0.805
Base+水平翻转	79.7	82.0	0.811
Base+垂直翻转	79.4	81.8	0.807
Base+模糊	79.8	82.3	0.812
Base+随机裁剪	79.8	82.1	0.812
Base+旋转	79.9	82.4	0.812
Base+涂装增广	80.10	82.70	0.813

由表可知,经过舱门涂装增广训练得到的 YOLOv5s 的准确率、召回率、*mAP* 分别达到了 82.9%、84.8% 和 0.837。相较于 5 种基础增广方法,舱门涂装增广的 YOLOv5s 的检测准确率、召回率、*mAP* 分别高了 3.0%、2.4% 以及 2.4%,提升值高于传统的基础增广方法。可以得出结论,相较于现有的基础增广方法,舱门图像涂装增广方法对 YOLOv5s 网络在复杂涂装场景下的检测准确率、召回率和 *mAP* 的提升更大。

综上所述,舱门涂装增广方法可以提升在复杂涂装舱门场景下的舱门检测网络的准确率、召回率与 *mAP* 三项检测指标,并且提升效果优于基础增广方法。

4 结束语

针对登机桥自动驾驶系统中存在的舱门检测网络对

复杂涂装舱门图像检测指标较低的问题,本文的研究贡献和创新点如下:针对舱门检测网络对复杂涂装舱门图像检测指标较低的问题,本文提出了舱门一边缘一舱门的方式重建舱门图像。实验证明,舱门涂装增广方法可以提升在复杂涂装舱门场景下的舱门检测网络的准确率、召回率与 *mAP* 三项检测指标,并且提升效果优于基础增广方法。

参考文献:

[1] MORENO-BAREA F J, JEREZ J M, FRANCO L. Data augmentation Meta-Classifer scheme for imbalanced data sets [C] //2022 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2022: 1392 - 1399.

[2] ZHONG Z, ZHENG L, KANG G, et al. Random erasing data augmentation [C] //Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (7): 13001 - 13008.

[3] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6023 - 6032.

[4] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization [J]. ArXiv Preprint ArXiv: 1710.09412, 2017.

[5] KONNO T, IWAZUME M. Icing on the cake: An easy and quick post-learnig method you can try after deep learning [J]. ArXiv Preprint ArXiv: 1807.06540, 2018.

[6] TERRANCE V, GRAHAM W T. Dataset augmentation in feature space [C] //Proceedings of the International Conference on Machine Learning (ICML), Workshop Track, 2017, 3.

[7] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2414 - 2423.

[8] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series [J]. The Handbook of Brain Theory and Neural Networks, 1995, 3361 (10): 1995.

[9] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution [C] //Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016: 694 - 711.

[10] ULYANOV D, LEBEDEV V, VEDALDI A, et al. Texture networks: Feed-forward synthesis of textures and stylized images [J]. ArXiv Preprint ArXiv: 1603.03417, 2016.

(下转第 202 页)