

基于改进多尺度知识蒸馏的电芯蓝膜缺陷异常检测

朱文博¹, 怀珍豪¹, 陈红光²

¹上海理工大学, 上海

²上海贝特威自动化科技有限公司, 上海

收稿日期: 2022年11月2日; 录用日期: 2022年11月30日; 发布日期: 2022年12月9日

摘要

针对新能源电芯工业生产过程中, 蓝膜缺陷样本数量少导致模型拟合难度大、缺陷目标小导致检测精度差、缺陷目标定位难的问题, 本文提出一种改进多尺度知识蒸馏的方法。在多尺度知识蒸馏骨干网络上结合残差注意力模块, 使网络聚焦图像中的局部信息, 提高对于小目标的检测准确性。

关键词

电芯蓝膜, 缺陷异常检测, 无监督学习, 知识蒸馏, 注意力机制

Abnormal Detection of Blue Membrane Defects in Cell Based on Improved Multi-Scale Knowledge Distillation

Wenbo Zhu¹, Zhenhao Huai¹, Hongguang Chen²

¹University of Shanghai for Science and Technology, Shanghai

²Shanghai Better Way for Automation Company, Shanghai

Received: Nov. 2nd, 2022; accepted: Nov. 30th, 2022; published: Dec. 9th, 2022

Abstract

In the process of industrial production of new energy cells, the number of blue membrane defect samples is small, which leads to difficult model fitting, poor detection accuracy and difficult defect target positioning. This paper proposes an improved multi-scale knowledge distillation method.

文章引用: 朱文博, 怀珍豪, 陈红光. 基于改进多尺度知识蒸馏的电芯蓝膜缺陷异常检测[J]. 软件工程与应用, 2022, 11(6): 1248-1254. DOI: 10.12677/sea.2022.116127

The residual attention module is combined with the multi-scale knowledge distillation backbone network to make the network focus on the local information in the image and improve the detection accuracy of small targets.

Keywords

Cell Blue Membrane, Defect Abnormal Detection, Unsupervised Learning, Knowledge Distillation, Attention Mechanism

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前国内主要采用人工目视法检测蓝膜缺陷,但是这种方式效率低且人工成本高。而随着机器视觉技术的发展,面对单一场景的检测往往可以取得优异的效果,但面对复杂多变的场景时缺乏泛化性[1]。采用深度学习技术来解决这些局限性是目前比较常用的方法。在电芯工业生产中,正常蓝膜样本数量远远大于有缺陷的异常样本,并且缺陷面积较小。基于深度学习的表面缺陷检测中,缺陷样本少、缺陷小导致无法获得准确的检测结果,通常难以满足工业上的需求,基于无监督学习的异常检测方法通常用来解决此类问题[2]。

在该领域国内外学者进行了一系列研究,Kang [3]采用深度去噪编码器,利用其误差重建以此检测绝缘子图像中的缺陷。LSA [4]在其潜在空间训练自回归模型,而 OCGAN [5]试图强制将异常输入重构为正常输入。基于 GAN 的方法,如 f-AnoGan [6]、GANomaly [7]在 AnoGan [8]的基础上添加额外的编码器以减少推理时间,在异常检测与定位上表现良好,但在单类缺陷检测上效果还有待提高。以上这些无监督学习方法适用于异常样本少的场景,但是很难同时满足异常检测和精确定位,且模型计算量大、实用性较差。

知识蒸馏采用预先训练的网络作为教师网络,将学习到的特征“蒸馏”到学生网络中,在保证精度的同时减少模型的计算量。Salehi [9]等提出一种多尺度知识蒸馏方法,通过将—个在 ImageNet 上预先训练的源网络上的各层知识,提取到一个结构相同但模型尺寸较小的克隆网络上,可以有效利用教师网络中间层信息,表现出良好的异常检测性能。但是该方法缺乏对图像局部信息的关注,因而在针对小目标检测上效果有待提高。

在小目标检测领域,DSSD [10]采用反卷积技术用于 SSD 的所有特征图上,目的是得到放大比例的特征图,但是由于是在所有特征图上应用反卷积,增加了模型的复杂度导致速度变慢。

针对以上问题本文提出一种改进多尺度知识蒸馏的方法,采用正常数据训练解决缺陷样本少的问题。通过在多尺度知识蒸馏骨干网络上结合残差注意力模块,使网络聚焦图像中的局部信息,提高对于小目标的检测准确性,解决小目标缺陷检测精度差、定位难的问题。相较于传统缺陷异常检测,本文的方法更容易捕获小的图像细节,获得精确的异常定位结果。

2. 数据准备和预处理

锂电池电芯出厂时会包裹—层蓝膜,可以起到对电芯的保护作用。本文电芯蓝膜图像通过工业线扫平台以及工业线扫相机对电芯扫描得出,实验室环境下蓝膜图像数据采集装置示意图如图 1 所示,工业现场扫描得到的原始图像大小为 $4000 \times 18,000$ 。由于图片过大影响模型正常学习,而且蓝膜表面缺陷面积相对于整张图像非常小,因此将其切割成 1024×1024 大小的图像。



Figure 1. Sketch map of blue membrane image data acquisition device

图 1. 蓝膜图像数据采集装置示意图

在生产过程部分蓝膜中会由于外力作用产生破损(a)或者粘附异物(b)如图 2 所示, 其中异物以白色异物居多, 少量的黑色异物以及其他颜色异物, 在 3.2 实验分析一节中给出了不同缺陷的定位效果图。缺陷会极大影响蓝膜质量导致电芯出现安全问题, 有必要对蓝膜图像进行缺陷检测, 本文采用异常检测的方法识别有缺陷的蓝膜样本, 因此将含有缺陷的蓝膜图像统一视为异常样本。

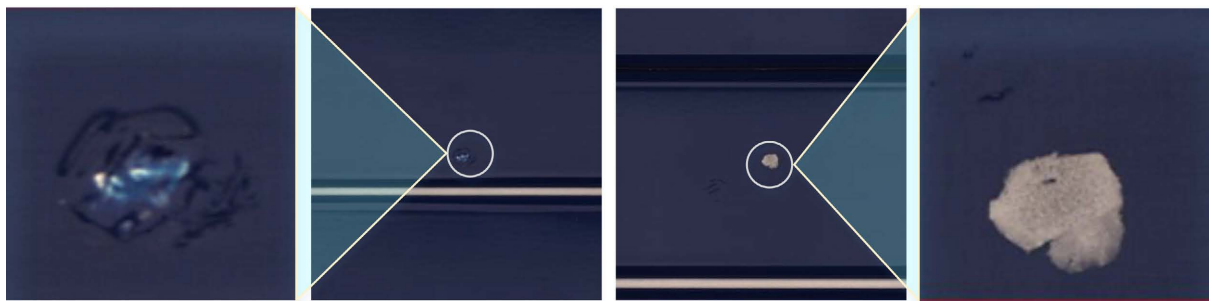


Figure 2. Blue membrane defect image

图 2. 蓝膜缺陷图

蓝膜数据集包含训练集与测试集, 训练集全部是没有缺陷的正常样本。实验结果通过梯度图可视化的方式进行展示。由于训练样本是正常图像数据, 因此只需对含有缺陷的测试集进行标注, 标注方法与图像分割相同, 标签类型为缺陷, 采用的标注工具为海康深度学习软件 VisionTrain1.4.1。训练集有 600 张正常图像, 测试集有 205 张图像, 包含 55 张正常图像, 150 张异常图像。

3. 改进的多尺度知识蒸馏

知识蒸馏是一种模型压缩的方法, 通过构建一个轻量化的小型网络, 利用性能更好的大模型的监督信息, 来训练这个小模型, 在降低计算成本的情况下, 能达到更好的性能和精度, 蒸馏得到的小模型在实际生活中更易于部署到算力受限的终端。在新能源电芯生产车间里由于环境限制, 蓝膜缺陷检测设备通常难以具备较大的计算能力, 因此知识蒸馏方法较为适用。

电芯蓝膜图像颜色表现为单一的蓝色, 背景像素统一无其他明显特征。基于多尺度知识蒸馏的异常检测方法在处理这类问题时表现出较好的效果。但是蓝膜图像中具有一部分尺寸细小的缺陷, 而多尺度知识蒸馏方法缺乏对图像局部信息的关注, 采用此方法处理小目标缺陷时往往检测率不高。因此考虑在该方法基础上进行改进, 以适应蓝膜缺陷检测的需求。本文通过在骨干网络上添加残差注意力机制增加网络对局部信息关注度, 以此提高小目标缺陷检测效果。

3.1. 多尺度知识蒸馏

多尺度知识蒸馏网络结构主要包括源网络 S 与克隆网络 C 两部分。通过训练一个较小的克隆网络模仿源网络在正常数据上的整个行为, 根据在训练中看到的正常样本, 识别出对模型来说新颖的测试时输入。由于克隆网络需要从正常数据的特征中预测样本的偏差, 所以需要正常数据特征极为了解, 训练时输入的训练集只包含正常数据, 目的是训练一个克隆网络 C 用于检测测试集中的异常图像。两个网络间行为的差异由一个总损失函数来表示。源网络采用 VGG-16 在一个大的自然数据集 ImageNet 上预训练而来, 克隆网络 C 与源网络 S 采用一样的结构, 但模型尺寸较小, 实验表明选择较小的克隆网络能有效提高网络性能[11]。

知识蒸馏为了节省计算成本致力于将知识从一个网络转移到另一个较小网络。但是多数知识蒸馏仅仅把 S 的最后输出传给 C , 这样做会丢失很多中间信息。由于神经网络各层对应着不同的抽象信息, 所以仅把最后一层输出传给 C , 只包含一部分信息。基于此, 采用一种多尺度知识蒸馏方法, 在网络几个关键层之后计算 S 与 C 的 loss 能传授给克隆网络更全面的信息。将知识视为激活函数的值, 在本文中知识定义为关键层激活值 a^{cp_i} 的距离 L_{val} 和方向 L_{dir} , CP_i 表示网络的第 i 个关键层。克隆网络 C 的激活值用 $a_c^{cp_i}$ 表示, 所以损失函数有两个, 距离损失为:

$$L_{val} = \sum_{i=1}^{N_{cp}} \frac{1}{N_i} \sum_{j=1}^{N_i} (a_s^{cp_i}(j) - a_c^{cp_i}(j))^2 \quad (1)$$

N_i 表示关键层 CP_i 的神经元数量, $a_s^{cp_i}(j)$ 表示关键层 CP_i 的第 j 个激活值。但是单一的距离损失难以完整表达向量间的相似性, 因此采用余弦相似度作为方向损失 L_{dir} 进行补充:

$$L_{dir} = 1 - \sum_i \frac{\text{vec}(a_s^{cp_i})^T \cdot \text{vec}(a_c^{cp_i})}{\|\text{vec}(a_s^{cp_i})\| \cdot \|\text{vec}(a_c^{cp_i})\|} \quad (2)$$

其中 $\text{vec}(x)$ 表示任意维矩阵 x 变换成一维向量的函数, 总损失 L_{total} :

$$L_{total} = \lambda L_{dir} + L_{val} \quad (3)$$

设置 λ 比例因子以调节两部分损失达到比例相同, 训练时使用 L_{total} 作为完全收敛标准。

3.2. 改进的多尺度知识蒸馏网络结构

Wang [12]等研究表明注意力机制可以让网络聚焦于图像中的一部分。为了提高小目标检测率, 在 VGG 骨干网络上, 额外添加注意力机制以创建不同分辨率的特征映射。本文将单阶段残差注意力模块置于第 2 层 conv2_2 以及第 7 层 conv7_3 之后。每个残差注意力模块结构由主干与掩码两个分支组成, 主干分支由两个残差块组成, 残差块包含 3 个卷积层。掩码分支经过一个上下采样与残差连接后输出注意特征图, 经 sigmoid 函数激活然后乘以主干分支的输出, 得出被注意特征图。最后经过残差块、relu 激活函数之后与克隆网络对应层的输出计算损失, 改进前后网络结构对比示意图如图 3 所示。

4. 实验结果及分析

4.1. 评价指标

在本文中, 评价指标采用 AUROC 曲线来反映模型的学习效果, 该曲线能够表现出模型的预测结果中正误的变化关系。曲线纵轴为真正类率(true positive rate, TPR),

$$TPR = \frac{TP}{TP + FN} \quad (4)$$

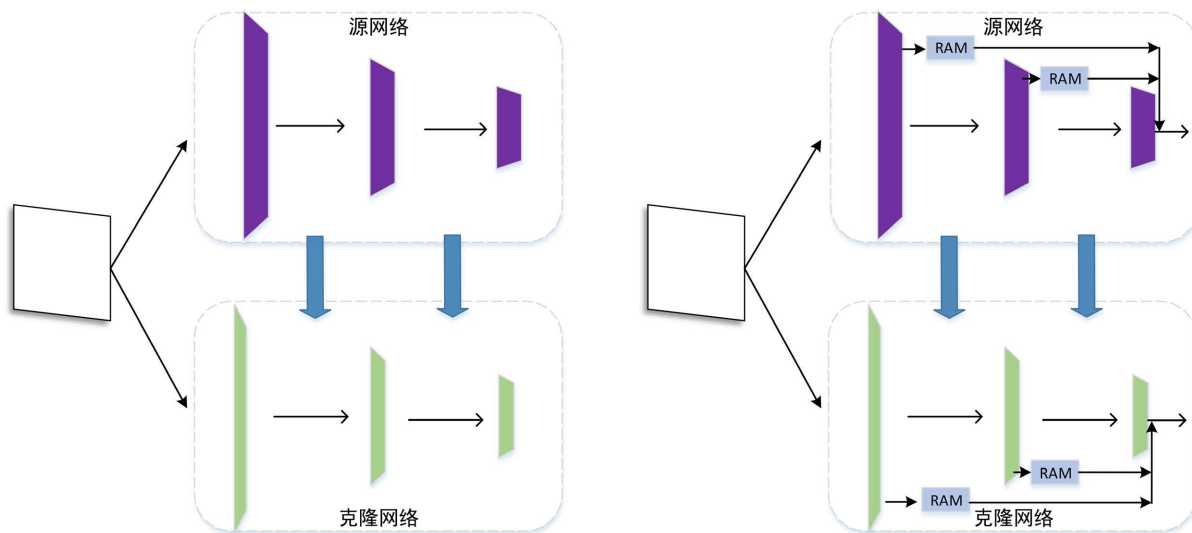


Figure 3. Improved multiresolution knowledge distillation structural comparison diagram
图 3. 改进多尺度知识蒸馏结构对比图

表示模型预测结果的正样本中实际正样本占有所有正样本的比例。 TPR 的值越大, 说明预测的正样本中实际的正样本越多。 TP 表示预测正确的正样本, FN 表示正样本被预测为负样本。曲线横轴为假正类率(false positive rate FPR),

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

表示模型预测的正样本中实际负样本占有所有负样本的比例。 FPR 的值越大, 说明预测的正样本中实际的负样本越多。 FP 表示负样本被预测成正样本, TN 表示正确预测的负样本。

4.2. 实验结果与分析

实验环境: Win10 操作系统, 处理器为 Intel(R) Core(TM) i5-11400H @2.70GHz, 显卡是 NVIDIA GTX 3080。采用 Pytorch 框架实现算法搭建, 训练 epoch 为 600, 学习率 0.001, 损失函数权重因子 λ 设置为 0.5。在训练过程中每 10 个 epoch 之后用验证集测试训练效果, 以此验证模型学习效果, 在经过 600 个 epoch 后, 训练 loss 值接近为零且趋于稳定状态。

本文在蓝膜数据集上与现有的几种方法 LSA、AnoGAN、VAE-grad、GT、CNN-Dict、MKD, 包括自监督、生成对抗网络、基于自编码器方法、基于多尺度知识蒸馏等进行了比较, 在公开数据集 MVTecAD 上选取了五种与蓝膜图像纹理相似的种类, 包括 Capsule、Carpet、Hazelnut、Wood、Leather, 最终的检测结果如表 1 与图 4 所示, 缺陷定位效果如图 5、图 6 所示。

Table 1. Detection AUROC (%) on datasets of blue membrane and MVTecAD
表 1. 在蓝膜以及 MVTecAD 数据集上的检测 AUROC (%)

Method	Small	Big	Capsule	Carpet	Hazelnut	Wood	Leather
LSA	71	87	71	74	80	75	70
AnoGAN	75.2	86.4	58	49	50	68	52
VAE-grad	81	83	86	67	74	89	71

Continued

GT	67.79	74.29	67.79	45.9	33	48.24	82.51
MKD	78	94	80.46	79.25	98.37	94.29	95.05
MKD + RAM (OURS)	81.91	95.66	78.53	82.73	93.15	95.33	95.87

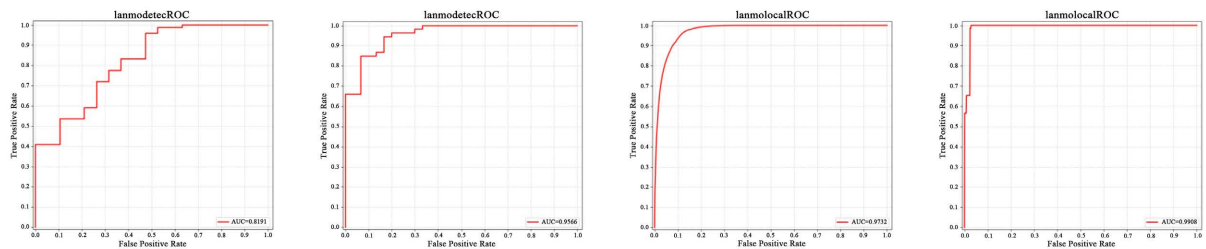


Figure 4. Auroc curve of detection on blue membrane dataset

图 4. 蓝膜数据集上检测 AUROC 曲线图

如图 5 所示是公开数据集 MVTecAD 中五个种类缺陷的异常定位测试效果图，图 6 是蓝膜缺陷图的单张定位测试效果图，用梯度图可视化的方式显示缺陷定位。其中 Wood、Carpet、Leather 相比其它两类与蓝膜纹理更为相似，因此检测效果更好，对于小目标缺陷的检测也更准确；而 Capsule 与 Hazelnut 的异常检测效果相对有待提高。所以本文方法对于图像背景像素单一、缺陷目标小场景下的异常检测更为适用。

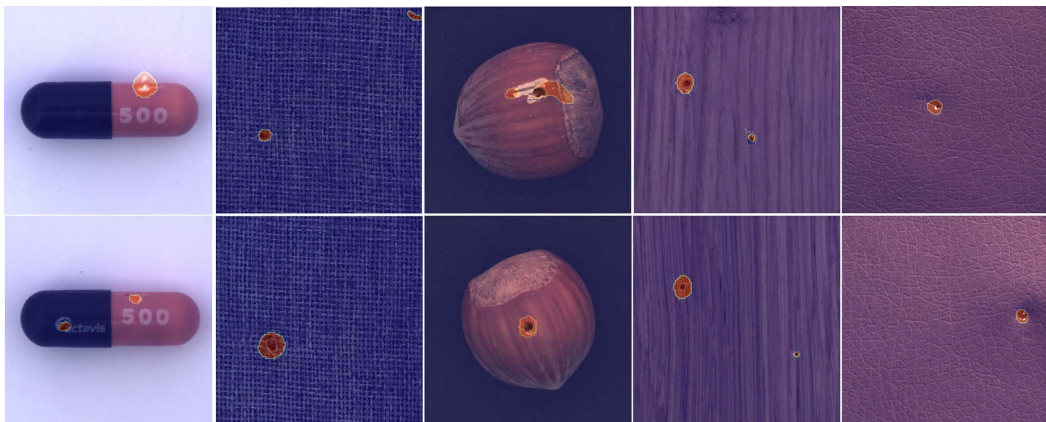


Figure 5. MVTecAD defect location rendering

图 5. MVTecAD 缺陷定位效果图

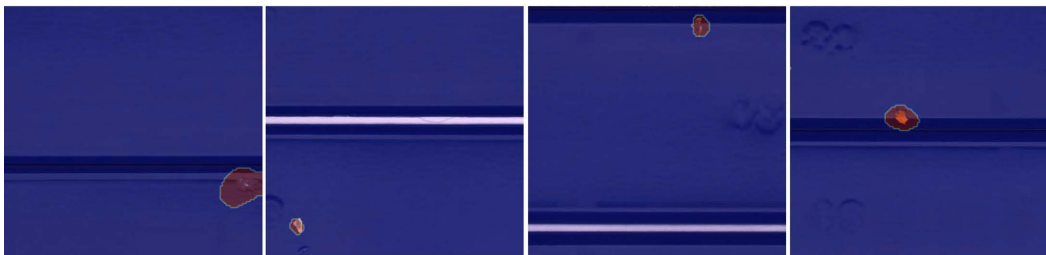


Figure 6. Blue membrane defect detection results

图 6. 蓝膜缺陷检测结果图

本文提出的在知识蒸馏骨干网络上添加残差注意力机制的方法,能够有效弥补传统知识蒸馏在检测小缺陷方面的不足,在蓝膜 Small 数据集上检测准确率较 MKD 方法提高了 3.91%,在其他场景的缺陷检测上准确率也有提高。而采用多尺度知识蒸馏方法进行训练,在节省资源的同时也让缺陷定位更为准确,因此无论是在异常检测任务上还是在异常定位任务上,在蓝膜数据集与公开数据集上本文提出的方法均表现出较好效果。

5. 总结

为了解决工业生产中电芯表面蓝膜缺陷样本量少导致模型拟合难度大、缺陷目标小导致检测精度差、定位难的问题,本文提出一种改进多尺度知识蒸馏的方法。在多尺度知识蒸馏骨干网络上结合残差注意力模块,使网络聚焦图像中的局部信息,提高对于小目标的检测准确性。采用无缺陷的正常数据进行训练,有缺陷的异常数据进行测试,在公开数据集 MVTecAD 和电芯蓝膜缺陷数据集上进行了异常检测以及异常定位实验。实验结果表明本文方法在处理缺陷背景呈相似像素的图像以及检测小目标缺陷方面的良好效果。本文方法的不足之处是面对公开数据集中背景变化较大的图像时,检测精度还有待提高。

参考文献

- [1] 陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报, 2021, 47(5): 1017-1034.
- [2] 吕承侃, 沈飞, 张正涛, 张峰. 图像异常检测研究现状综述[J]. 自动化学报, 2022, 48(6): 1402-1428.
- [3] Kang, G., Gao, S., Yu, L., et al. (2019) Deep Architecture for High-Speed Railway Insulator Surface Defect Detection: Denoising Autoencoder With Multitask Learning. *IEEE Transactions on Instrumentation and Measurement*, **68**, 2679-2690. <https://doi.org/10.1109/TIM.2018.2868490>
- [4] Abati, D., Porrello, A., Calderara, S., et al. (2019) Latent Space Autoregression for Novelty Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 481-490. <https://doi.org/10.1109/CVPR.2019.00057>
- [5] Perera, P., Nallapati, R. and Xiang, B. (2019) Ocgan: One-Class Novelty Detection Using Gans with Constrained Latent Representations. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 2898-2906. <https://doi.org/10.1109/CVPR.2019.00301>
- [6] Salehi, M., Arya, A., Pajoum, B., et al. (2021) Arae: Adversarially Robust Training of Autoencoders Improves Novelty Detection. *Neural Networks*, **144**, 726-736. <https://doi.org/10.1016/j.neunet.2021.09.014>
- [7] Schlegl, T., Seeböck, P., Waldstein, S.M., et al. (2017) Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. In: *Information Processing in Medical Imaging*, Springer, Cham, 146-157. https://doi.org/10.1007/978-3-319-59050-9_12
- [8] Schlegl, T., Seeböck, P., Waldstein, S.M., et al. (2019) f-AnoGAN: Fast Unsupervised Anomaly Detection with Generative Adversarial Networks. *Medical Image Analysis*, **54**, 30-44. <https://doi.org/10.1016/j.media.2019.01.010>
- [9] Salehi, M., Sadjadi, N., Baselizadeh, S., et al. (2021) Multiresolution Knowledge Distillation for Anomaly Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, 20-25 June 2021, 14902-14912. <https://doi.org/10.1109/CVPR46437.2021.01466>
- [10] Fu, C.Y., Liu, W., Ranga, A., et al. (2017) Dssd: Deconvolutional Single Shot Detector. arXiv preprint arXiv:1701.06659, 2017.
- [11] Bergmann, P., Fauser, M., Sattlegger, D., et al. (2019) MVTec AD—A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 9592-9600. <https://doi.org/10.1109/CVPR.2019.00982>
- [12] Wang, F., Jiang, M., Qian, C., et al. (2017) Residual Attention Network for Image Classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 3156-3164. <https://doi.org/10.1109/CVPR.2017.683>