

一种基于电致发光图像的无监督光伏电池异常检测方案

田家源¹, 王玉峰²

(1. 南京邮电大学波特兰学院, 江苏 南京 邮编210023; 2. 南京邮电大学通信与信息工程学院, 江苏 南京 邮编210023)

摘要: 电致发光 (EL) 成像技术结合基于特征提取的监督学习方法常常被用于光伏电池的异常检测, 但监督学习不仅标记成本高昂, 且难以检测出训练集中没有的异常模式。无监督学习方法有助于解决上述问题, 然而如何高效利用多尺度特征进行无监督异常检测仍缺乏成熟的范式。对此, 本文提出了一种新颖的基于 EL 图像的无监督光伏电池异常检测方案—MFRAD。首先, 利用深度卷积网络提取光伏电池 EL 图像的多尺度特征。其次, 针对高维多尺度特征设计了高效的对抗重构模块。最后, 综合特征空间和潜在空间中的重构误差组成异常评分, 实现光伏电池异常的有效检测。实验结果表明, MFRAD 在单晶和多晶光伏电池 EL 数据集上各取得了 0.956, 0.868 的 ROC-AUC, 优于其他无监督异常检测方法。

关键词: 电致发光; 无监督学习; 深度卷积网络; 多尺度特征; 对抗性重构

中图分类号: TP391.41

文献标志码: A

doi: 10.11959/j.issn.2096-3750.XXXX.

An Unsupervised Photovoltaic Module Anomaly Detection Scheme Based on EL Images

TIAN Jiayuan¹, WANG Yufeng²

1. Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210023, China

2. College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu 210023, China

Abstract: Electroluminescence (EL) imaging combined with feature-based supervised learning was often used for photovoltaic (PV) cell anomaly detection. However, supervised methods incurred high labeling costs and failed to detect anomaly patterns not present in the training set. Unsupervised learning was employed to address these issues. Nevertheless, no established paradigm had existed for the efficient utilization of multi-scale features in unsupervised anomaly detection. To tackle this challenge, a novel unsupervised PV cell anomaly detection method, MFRAD, was proposed. Multi-scale features of EL images were first extracted using a deep convolutional network. Then, an efficient adversarial reconstruction module was designed for high-dimensional multi-scale features. Finally, reconstruction errors in both the feature space and latent space were combined to generate anomaly scores, enabling effective detection of PV cell anomalies. The experimental results showed that MFRAD achieved ROC-AUCs of 0.956 and 0.868 on monocrystalline and polycrystalline photovoltaic cell EL datasets, outperforming other unsupervised anomaly detection methods.

Key words: electroluminescence (EL), unsupervised learning, deep convolutional network (DCN), multi-scale features, adversarial reconstruction

0 引言

国际能源署最新报告[1]显示, 到2050年, 可

再生能源将占近一半的全球电力供应。光伏 (PV) 发电通过光伏电池将太阳辐射转化为电能, 具有安全、可再生和清洁等优势。然而, 光伏电池在制造过程中可能受焊接缺陷以及晶体生长缺陷的影响,

收稿日期: XXXX-XX-XX; 修回日期: XXXX-XX-XX

运行中还可能因树枝掉落、冰雹或热应力产生机械损伤[2]。这些异常不仅会降低光伏电池的发电效率，还会缩短其使用寿命，造成重大经济损失。因此，对光伏电池进行有效的异常检测，是保障太阳能电站发电量的关键。

电致发光（EL）成像是一种广泛用于光伏电池异常检测的无损技术[3][4]，可以揭示光伏电池中的微裂纹、指状断裂、污染和焊接缺陷等异常。然而，人工检查EL图像不仅耗时，而且需要专业知识，因此基于EL图像的光伏电池自动异常检测是一种高效且经济的解决方案。

随着深度学习技术的快速发展，卷积神经网络（CNN）已成为各种计算机视觉任务的强大工具，包括EL图像的异常检测。例如，文献[5]提出先通过CNN提取图像特征，再使用支持向量机等传统机器学习方法对缺陷光伏电池进行分类；文献[6]针对光伏电池设计了改进的YOLO模型，用于光伏电池的快速异常检测；基于自注意力机制捕捉长距离依赖的能力，Transformer的变体也被用于光伏电池异常检测[7][8]；文献[9]利用生成对抗网络（GAN）[10]扩充有限且不平衡的EL数据集，再利用CNN实现缺陷检测。

上述工作遵循监督学习范式，即使用大量标记样本，将光伏电池异常检测构建为监督分类问题，并通过交叉熵损失训练各种CNN模型，以对EL图像中的不同类型光伏电池故障进行分类。尽管监督学习在某些场景下效果良好，但存在以下不足。首先，标记足够的异常样本不仅耗时，还需要领域专业知识。此外，光伏电池的监督异常检测通常无法有效检测未知类型的异常，进一步降低了其实用价值。

而无监督学习方法则不依赖异常EL图像标注，旨在学习EL图像的正常模式并为每张EL图像推断异常分数，这有助于节省人力成本并可以检测出未知的异常。一种广泛使用的学习正常数据分布的方法是图像重建，即给定正常图像 x ，尝试在潜在空间中构建其潜在表示 z ，并使 z 相应的重建实例 $G(z)$ 与 x 尽可能相似。由于潜在空间仅学习了训练数据集中正常数据的潜在分布，异常样本与其重构的相似度会低于正常样本。遵循这一范式，深度自编码器[11]被用于光伏电池的异常检测。类似地，文献[12]结合长短时特征检测光伏电池异常，其

中，通过去噪自编码器从输入的光伏电池图像中提取短时特征，利用在ImageNet数据集上预训练的深度CNN提取光伏电池图像的长时特征作为先验知识以提升检测性能。近年来，GAN通过生成器和判别器的博弈，展现出训练鲁棒生成模块的能力，并在无监督异常检测任务中显示出潜力。例如，f-AnoGan[13]提出了一种基于GAN的快速异常检测方法，其中训练编码器将每个图像 x 快速映射到其学习的潜在表示 z ，通过融合判别器特征残差和图像重建误差的异常分数检测异常。

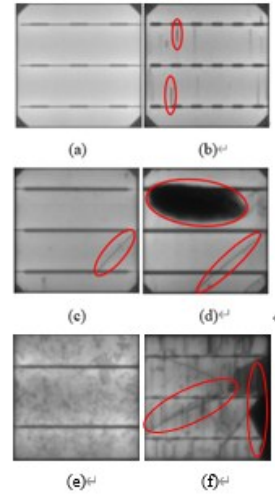


图1 光伏电池电致发光图像:(a)单晶正常;(b)~(d)单晶缺陷;(e)多晶正常;(f)多晶缺陷。

如图1所示，光伏电池的异常具有以下特点，这些特点使得基于图像级重建的异常检测方法效果不佳。

低对比度和异常微小：如图1所示，由于EL图像为灰度图像，所以正常区域与异常区域之间的对比度较低。例如，图1（b）中的指状断裂和图1（c）、（f）中的微裂纹等小异常，容易被光伏电池的背景所掩盖，增加了检测难度。

复杂的异常模式：如图1（d）所示，微裂纹通常呈细长形，而黑芯则通常为大面积不规则区域。因此，同时有效检测这些复杂的异常模式具有挑战性。

为应对这些挑战，本文提出了基于光伏电池EL图像的无监督异常检测框架—MFRAD。不同于像素级图像重建，MFRAD旨在提取、融合和重建深度多尺度特征，并在特征空间中检测异常。具体而言，本工作的贡献总结如下：

在MFRAD框架中，我们利用预训练的深度卷

积网络提取光伏电池EL图像的多尺度特征，以有效捕捉光伏电池丰富的局部细节和空间上下文信息。在此基础上，我们在特征空间进行异常检测，有效应对了光伏电池EL图像中低对比度和复杂的缺陷模式带来的挑战。

我们针对高维多尺度特征设计了一个高效的对抗性重构模块，为基于多尺度特征提取的无监督异常检测提供了一种新的思路。具体来说，该模块采用编码器—解码器—编码器结构，在保留多尺度特征空间分辨率的同时仅对通道维度进行压缩、重构与再压缩，从而在特征空间和潜在空间施加一致性约束，实现了对高维多尺度特征的高效学习。同时，我们引入了基于 Wasserstein^[14]距离的对抗训练策略，有效解决了传统GAN在对抗重构中常见的训练不稳定与收敛困难问题。最终，我们将多尺度特征空间与潜在空间的重建误差进行融合，并以其作为异常分数，从而实现了光伏电池的高效异常检测。

我们在真实光伏电池图像数据集ELPV (<https://github.com/zae-bayern/elpv-dataset>) 上进行了广泛的实验评估，该数据集包含了单晶和多晶光伏电池EL图像，可以覆盖实际应用中光伏电池EL图像异常检测的绝大部分场景。结果表明，所提出的方法优于其他无监督异常检测方案。此外，我们还进行了消融实验以验证对抗重建模块中关键组件的有效性。

本文结构如下：第1节回顾了异常检测的相关工作；第2节详细介绍了所提出的方法MFRAD；第3节讨论了实验结果，包括对比实验和消融研究；最后，在第4节总结本文并概述未来工作。

1 相关工作

1.1 图像级重建方案

自编码器(AE)[15][16]通常用于图像级重建。这类模型仅使用正常图像训练，并计算图像与其重建结果的像素级距离，如 L_2 距离[17]和结构相似性指标SSIM[18]。为高质量重建正常图像，文献[19]将编码器各层提取的信息通过跳跃连接直接传递至解码器对应层，使解码器能更好地重建输入图像的结构和纹理细节。

此外，对抗训练策略[20]也被引入该领域。文献[21]将生成对抗网络(GAN)与卷积自编码器

(CAE)结合，通过潜在空间的额外重建约束及判别器的辅助，提升了正常图像的重建质量。但GAN本身存在训练不稳定和收敛困难的问题。对此，文献[13]融合了 Wasserstein生成对抗网络(WGAN)与自编码器，Wasserstein距离的引入为训练提供了更有效且连续的梯度，显著提升了训练的稳定性和重建质量。

然而，上述基于图像重建的异常检测方法旨在最小化原始图像的重建误差，这不可避免地促使模型主要关注像素级信息，从本质上限制了其全面捕捉全局特征的能力，进而降低了推理阶段的异常检测性能。

1.2 特征提取方案

为解决上述问题，基于深度特征的方法从正常样本中提取具有表现力的多尺度特征，进而通过特征分布的差异实现异常检测。文献[22]，[23]采用反向蒸馏技术，学生网络以教师模型的单类嵌入特征为输入，尝试重建教师的多尺度表示，重建误差越大，异常的可能性越高；文献[24]通过聚类算法学习预训练深度CNN所提取特征中的类别信息，并据此实现多类别异常检测；文献[25]将多尺度特征提取与条件归一化流建模相结合，确保实时高效的异常检测；文献[26]将图像分割为局部块，并采用高斯分布对从这些局部块中提取的多尺度特征进行建模；另有研究使用预训练的CNN提取图像的多尺度特征，通过CAE计算特征重建误差来区分正常与异常样本^{[27][28]}。然而，对于如何有效且高效地利用高维的多尺度特征进行异常检测，目前仍缺乏一个统一而成熟的范式。

2 光伏电池异常检测框架—MFRAD

图2展示了我们提出的无监督光伏电池异常检测框架MFRAD，该框架包含两个核心模块：多尺度深度特征提取与融合模块，以及对抗性重构模块。对于输入的光伏电池EL图像 x ，我们首先利用预训练的深度卷积神经网络提取不同尺度的特征图，并将这些多尺度特征图融合为一个统一的多尺度特征表示。在此基础上，我们提出了对抗性重建模块，由“编码器(E_1)—解码器(D_c)”结构的生成器、判别器(D)以及一个附加编码器(E_2)组成，以高效最小化所获融合多尺度特征表示与其潜在表示之间的重建误差。在推理阶段，通过计算

测试图像的多尺度深度特征及其潜在表示的双重重构误差，生成异常分数以实现异常检测。

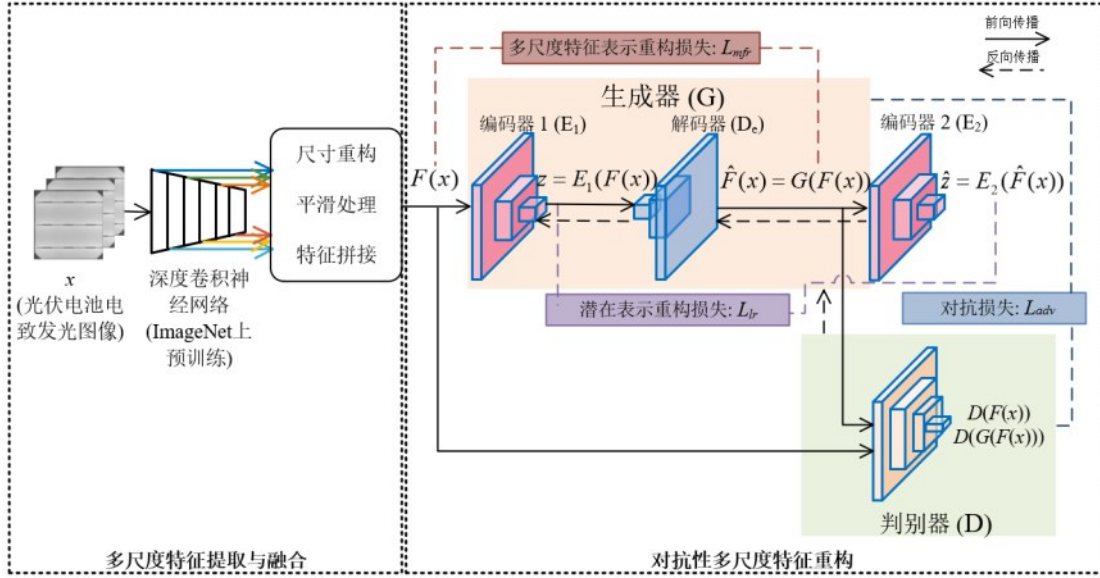


图2 光伏电池无监督异常检测方法MFRAD的框架图

2.1 多尺度特征提取与融合

文献[29]中的研究证明，与原始图像相比，在特征空间中能更有效地区分正常与异常样本。因此，我们首先采用在 ImageNet 预训练的参数冻结深度 CNN 模型，从光伏电池图像中提取多尺度特征，用于在特征空间中进行异常检测。本工作选用 VGG-19[30]作为特征提取器，其包含 16 个卷积层，每层可基于不同感受野输出特征图。在这里其他预训练的深度 CNN 模型同样适用。

假设输入光伏电池图像 x 的宽度为 w ，高度为 h ，通道数为 c ，当 x 通过预训练的 VGG-19 网络时，可以获得特征图集合 $\{f_1(x), f_2(x), \dots, f_l(x)\}$ 。其中，第 l 个卷积层输出的特征图 $f_l(x)$ 的大小为 $h_l \times w_l \times c_l$ 。因为 VGG-19 的各个卷积层的感受野是逐级递增的，所以每个卷积层学习到的光伏电池图像的特征各不相同，较小的感受野能学习到光伏电池纹理结构等低层次特征，而较大的会学习到更多的全局的或更高层次的信息。这些特征图共同构成了初步的光伏电池 EL 图像的多尺度特征表示，兼具局部细节与多层次空间上下文信息。

由于不同卷积层输出的特征图的空间大小不同，需要通过处理将各个尺度特征图融合为统一表示，以便于后续处理。所以，如图3所示，我们首先在保持通道数 c_l 不变的情况下将所有特征图上采样调整到与原始图像相同的宽度和高度 ($h \times w$)：

$$\bar{f}_l(x) = \text{resize}(f_l(x)) \quad (1)$$

其中， $\bar{f}_l(x)$ 为尺寸重构的特征图，大小为 $h \times w \times c_l$ 。

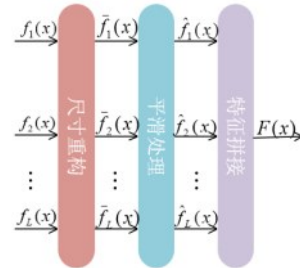


图3 多尺度特征的具体融合步骤

接着，对调整过的特征图进行平滑操作，具体来说，是用一个平均滤波器在对齐的特征图 $\bar{f}_l(x)$ 上按一定步长进行卷积：

$$\hat{f}_l(x) = \text{smooth}(\bar{f}_l(x)) \quad (2)$$

其中， $\hat{f}_l(x)$ 为经过平滑处理的特征图，大小为 $h_o \times w_o \times c_l$ 。

该操作可以平滑特征图上特征的变化，使特征对光伏电池图像上的污渍、灰尘等噪声有一定的鲁棒性，并降低计算复杂度。然而，滤波器过小会导致平滑效果不足，无法有效减少噪声对特征表示的影响。而滤波器过大，则会降低特征的连续性，模糊重要的特征。因此，实验中滤波器采取大小为 4×4 ，步长为 4 的折中方案。最后，将经过平滑操

作得到的所有特征图连接起来：

$$F(x) = \text{cat}(\hat{f}_1(x), \hat{f}_2(x), \dots, \hat{f}_l(x)) \quad (3)$$

其中， $F(x)$ 为拼接后的多尺度特征表示，大小为

$$h_o \times w_o \times c_o, o = \sum_{l=1}^L c_l$$

经过上面的操作，我们最终得到了一个完整的，因含有光伏电池EL图像丰富的局部细节和空间上下文信息而具有良好判别性的，同时又方便于后续异常检测任务的高维多尺度特征表示 $F(x)$ 。

2.2 对抗性多尺度特征重构

对于最终得到的多尺度特征表示 $F(x)$ ，我们需利用其鉴别性以实现异常检测。针对这个特征维度很大的多尺度特征表示，我们设计了一个基于Wasserstein距离的对抗性特征重构模块，它由生成器，判别器和一个附加编码器组成。

2.2.1 生成器和附加编码器

我们的生成器选用“编码器—解码器”的重构结构，编码器和解码器由对称的三个卷积层组成，每个卷积层实现了 1×1 的卷积和ReLU操作，从而在保持空间分辨率的同时压缩通道深度并进行后续重构。需要注意的是，虽然只作用于通道维度，但结合实验结果，这一设计可以有效且高效学习多尺度特征表示。具体来说，生成器中的编码器 E_1 将光伏电池EL图像的高维多尺度特征表示 $F(x)$ 压缩为潜在表示 $z = E_1(F(x))$ 。随后解码器 D_e 将 z 重构为 $\hat{F}(x) = G(F(x)) = D_e(E_1(F(x)))$ 。其中， $G(\cdot) = D_e(E_1(\cdot))$ ，重构的多尺度特征 $\hat{F}(x)$ 被判别器视为“假样本”。接着，和 E_1 结构完全相同且参数共享的编码器 E_2 对重构后的多尺度特征 $\hat{F}(x)$ 进行二次压缩，得到重构的潜在表示 $\hat{z} = E_2(\hat{F}(x)) = E_2(G(F(x)))$ 。我们通过联合特征空间中 $F(x)$ 与 $\hat{F}(x)$ 的重构误差，以及潜在空间中 z 与 \hat{z} 的重构误差来评估重建质量。这种双重编码机制显著增强了网络对光伏电池EL图像关键特征的捕捉能力。由于训练时仅使用正常样本，异常样本在特征空间或潜在空间中的重构误差会显著高于正常样本，从而实现异常检测。

2.2.2 判别器

我们的判别器引用了WGAN-GP[14]的思想。此时，原始的多尺度特征表示 $F(x)$ 被视作真实的样本，生成器输出的重构的多尺度特征表示

$\hat{F}(x) = G(F(x))$ 被视作虚假的样本。二者同时输入判别器，分别得到判别分数 $D(F(x))$ 和 $D(\hat{F}(x))$ 。判别器由与编码器相同的卷积层和一个全连接层组成。与传统GAN的判别器不同，我们判别器的最后一层全连接层不使用Sigmoid激活函数。此外，我们通过引入Wasserstein距离和梯度惩罚来修改生成器和判别器的损失函数。Wasserstein距离可以有效衡量两个概率分布之间的差异，并适用于两个分布不重叠的情况，从而减轻了传统GAN的不稳定和收敛问题。梯度惩罚则用于平滑判别器的梯度，使其满足1-Lipschitz约束，从而保证Wasserstein距离的有效计算和稳定的训练。

2.3 损失函数

2.3.1 生成器损失

生成器的损失函数 L_G 如公式(4)由 L_{mfr} ， L_{lr} ， L_{adv} 按权重 w_{mfr} ， w_{lr} ， w_{adv} 相加得到：

$$L_G = w_{mfr}L_{mfr} + w_{lr}L_{lr} + w_{adv}L_{adv} \quad (4)$$

其中，对抗损失 L_{adv} 和潜在表示重构损失 L_{lr} 被设计用于辅助模型学习正常多尺度特征，而异常检测的核心贡献仍来自多尺度特征表示的重构误差 L_{mfr} 。因此经验上 w_{mfr} 应较大，辅助权重 w_{adv} 和 w_{lr} 应较小。经实验验证，将权重设置为 $w_{mfr}:w_{lr}:w_{adv} = 10:1:1$ 时效果最佳。当多尺度特征重构损失权重过大时，模型会过度专注于重构任务，两个辅助损失的作用几乎失效，造成检测性能下降；而权重过小，则会造成多尺度特征重构的贡献不足，异常检测性能同样会下降。

多尺度特征表示重构损失 L_{mfr}

L_{mfr} 为多尺度特征表示的重构损失，用于衡量特征空间中多尺度特征表示 $F(x)$ 与其重构 $\hat{F}(x) = G(F(x))$ 之间的差异。我们希望误差尽可能小，即生成器能够尽可能准确地重构正常样本的多尺度特征。对于从输入数据分布 p_x 中抽取的给定输入 x ，重构损失定义为 x 的多尺度特征表示 $F(x)$ 和其重构 $\hat{F}(x)$ 之间的 l_2 距离，如公式(5)所示：

$$L_{rec} = E_{x \sim p_x} \|F(x) - G(F(x))\|_2 \quad (5)$$

潜在表示重构损失 L_{lr}

L_{lr} 为双重编码损失，是潜在表示 z 和 \hat{z} 之间的一致性约束。通过减小潜在空间中 $z = E_1(F(x))$ 和 $\hat{z} = E_2(G(F(x)))$ 之间的误差，可以辅助提高生成器对正常样本的多尺度特征表示的重构效果。编码

损失定义为 z 和 \hat{z} 之间 l_2 距离:

$$L_{lr} = E_{x \sim p_s} \left\| E_1(F(x)) - E_2(G(F(x))) \right\|_2 \quad (6)$$

对抗损失 L_{adv}

L_{adv} 为对抗损失, 生成器的目的是生成与真实多尺度特征 $F(x)$ 难以区分的样本 $\hat{F}(x) = G(F(x))$ 以欺骗判别器。因此, 生成器试图最大化 $D(G(F(x)))$, 所以生成器的对抗损失如公式 (7) 所示:

$$L_{adv} = -E_{x \sim p_s} [D(G(F(x)))] \quad (7)$$

2.3.2 判别器损失

判别器的损失函数 L_D 由两个部分组成: 基于 Wasserstein 距离的判别器损失 L_{WGAN} 和梯度惩罚项 L_{gp} 。如公式 (8) 所示:

$$L_D = L_{WGAN} + \lambda L_{gp} \quad (8)$$

其中, λ 是惩罚系数, 用于控制梯度惩罚的强度。

$$L_{WGAN} = -L_{adv} - E_{x \sim p_s} [D(F(x))] \quad (9)$$

其中, 第一项 L_{WGAN} 用于确保判别器将给真实样本 $F(x)$ 高分 $D(F(x))$, 而给假样本 $\hat{F}(x) = G(F(x))$ 低分 $D(G(F(x)))$ 。注意, 该项中 L_{adv} 与公式 (7) 相同。

第二项为梯度惩罚项, 旨在保证判别器梯度满足 1-Lipschitz 约束以有效计算真假样本分布之间的 Wasserstein 距离, 并稳定训练过程。梯度惩罚项如公式 (10) 所示:

$$L_{gp} = E_{x \sim p_s} \left[\left(\left\| \nabla_{\tilde{F}(x)} D(\tilde{F}(x)) \right\|_2 - 1 \right)^2 \right] \quad (10)$$

其中, $\tilde{F}(x)$ 是在真实样本 $F(x)$ 和 $\hat{F}(x)$ 间均匀采样得到。

2.4 异常评分

在测试阶段, 判别器不参与工作, 因为异常分数仅基于重构误差计算。对于输入的光伏电池图像 x , 异常分数综合了特征空间中的重构误差和潜在空间中的编码器误差。如公式 (11) 所示, 异常分数由 $F(x)$ 和 $\hat{F}(x)$ 间的 l_2 距离以及 z 和 \hat{z} 之间 l_2 距离按权重 w_{mfr} , w_{lr} 相加得到:

$$A(x) = w_{mfr} L_{mfr} + w_{lr} L_{lr} \quad (11)$$

其中 $w_{mfr}:w_{lr}$ 仍为 10:1, 多尺度特征重构误差 L_{mfr} 为主要异常评分来源, 同时兼顾潜在表示重构误差 L_{lr} 的辅助效果。

接着我们按公式 (12) 将异常分数限制到 [0,1]

的范围内:

$$A_s(x_i) = \frac{A(x_i) - \min(T)}{\max(T) - \min(T)} \quad (12)$$

其中, $T = \{T_i = A(x_i), x_i \in D_{test}\}$, 即所有异常评分的集合, $A_s(x_i)$ 为归一化的异常得分。

如果异常分数大于预定的阈值 T , 则光伏电池图像被分类为异常样本。阈值 T 的设置见下文。

3 实验

在本节中, 我们对真实数据集进行了全面实验, 以证明我们的方案优于现有的无监督异常检测方案。此外, 我们还设计了消融研究, 以探究我们提出的 MFRAD 框架中关键组件的效果。

3.1 实验数据和设置

本研究采用一个真实的光伏电池数据集对所提框架进行评估。该光伏电池数据集为 300×300 像素的 8 位灰度图像, 包含单晶光伏电池 (588 个正常样本, 120 个异常样本) 和多晶光伏电池 (920 个正常样本, 134 个异常样本)。正常光伏电池图像用于训练, 并按 4:1 的比例分为训练集和测试集。为了克服训练样本不足的问题并提高鲁棒性, 我们采用了在线数据增强技术。具体来说, 每个批次中, 50% 的图像通过随机水平翻转和旋转 (-15° 到 15°) 进行了增强, 另外 50% 则通过颜色抖动 (随机改变图像的亮度和对比度) 进行了增强。

实验基于 Pytorch 1.2.0 和 Python 3.7 实现, 所有模型都在 NVIDIA RTX 2060 GPU 中进行训练和测试。实验中, 学习率设置为 $2e-4$, 使用 Adam 作为优化器, Adam 的参数设置为 (0.5, 0.999)。模型进行了 300 次迭代训练, 批量大小为 4。具体的, 为了方便计算, 输入图像的分辨率设置为 256×256 像素。在多尺度特征提取与融合模块中, 特征层数 L 经验性地设置为 16, 即选取所有特征层。尺寸重构中上采样的方式设置为最近邻插值。平滑操作中, 平均滤波器的大小为 4×4 , 步长为 4。对于对抗性重构模块, 自编码器潜在空间维度的选择基于多尺度特征 ($c_o = 5504$) 的主成分分析, 通过保留累计方差达到 90% 的主成分, 得到降维后的维度为 222, 因此潜在空间的维度设置为 200。 $w_{mfr}:w_{adv}:w_{lr}$ 设置为 10:1:1。惩罚系数 λ 经验性地设置为 10。选择阈值 T 以使 Youden 指数 J 最大化:

$$J = TPR - FPR \quad (13)$$

其中，TPR为真阳性率，FPR为假阳性率，分别在以下方程式（16）和（18）中定义。

3.2 基准方案和性能指标

我们将MFRAD和以下三种无监督图像异常检测方案进行对比：

GANomaly^[21]：采用编码器-解码-编码器的结构，通过最小化正常样本的图像级重构误差和潜在向量级重构误差，并以判别器辅助以提高正常样本的重构质量。假设只用正常样本训练的模型无法同样高质量的重构异常图像，则在测试阶段异常图像的重构误差会高于正常图像，从而得到异常分数来判定图像是否异常。

f-AnoGAN^[13]：采用两阶段训练方法，首先在正常样本上训练WGAN，然后冻结其参数，用WGAN的生成器作为解码器来训练编码器。该训练旨在最小化正常图像的重构误差及其潜在表示，以提高重构质量。假设仅在正常样本上训练的网络对于异常图像会有更高的重构误差，从而根据重构误差计算分类的异常分数。

DFR^[27]：利用预训练的深度神经卷积网络提取正常图像的多尺度特征，再利用卷积自动编码器对处理后的多尺度特征进行重构。假设只用正常样本训练的网络无法同样高质量的重构异常图像的多尺度特征表示，则在测试阶段异常图像多尺度特征的重构误差会高于正常图像，从而得到异常分数来判定图像是否异常。

DACL^[31]：本工作将文献[31]的方案简称为DACL，其同样采用对抗重构的策略。在重建阶段，同时采用信道注意力和像素注意力。此外，提出了一种一致性损失函数作为生成网络的优化目标，该损失由像素差异，结构相似度度量和梯度信息组成，以利用图像多种模态之间的差异来提高异常检测的性能。

我们采用受试者工作特征曲线下的面积（ROC-AUC）和精确率-召回率曲线下面积（PR-AUC）作为二分类模型性能的主要衡量指标。此外，我们还考虑了四个其他指标来评估性能。它们分别是准确率（Accuracy）、精确率（Precision）、召回率（Recall）和F1分数（F1-Score）。其中，F1分数取决于真阳性（TP）：异常样本被正确分类为异常的情况；真阴性（TN）：正常样本被正确分类为正常的情况；假阳性（FP）：正常样本被错误地

分类为异常的情况；以及假阴性（FN）：异常样本被错误地分类为正常的情况。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

$$Precision = \frac{TP}{TP + FP} \quad (15)$$

$$Recall = \frac{TP}{TP + FN} \quad (16)$$

$$F1 - score = 2 \cdot \frac{Precision \times Recall}{Precision + Recall} \quad (17)$$

其中，公式（16）给出的召回率指标也被称为真阳性比率（TPR）或检出率。

表1 单晶光伏电池数据集上的对比结果

模型	ROC-AUC	PR-AUC	准确率	精确率	召回率	F1-分数
GANomaly	0.881	0.900	0.817	0.818	0.858	0.837
f-AnoGAN	0.926	0.920	0.870	0.941	0.792	0.860
DFR	0.940	0.951	0.882	0.877	0.892	0.884
DACL	0.936	0.949	0.854	0.951	0.808	0.874
MFRAD	0.956	0.965	0.899	0.887	0.917	0.902

表2 多晶光伏电池数据集上的对比结果

模型	ROC-AUC	PR-AUC	准确率	精确率	召回率	F1-分数
GANomaly	0.771	0.797	0.735	0.765	0.679	0.719
f-AnoGAN	0.803	0.838	0.743	0.933	0.522	0.670
DFR	0.839	0.877	0.799	0.864	0.710	0.779
DACL	0.831	0.860	0.781	0.890	0.698	0.782
MFRAD	0.868	0.899	0.810	0.888	0.709	0.788

3.3 实验结果

表1和表2给出了MFRAD和所有比较方案在数据集上的实验结果，最佳性能值加粗显示。表3列出了各个方法的推理速度FPS指标。

表3 推理速度FPS指标

模型	FPS(张/秒)
GANomaly	61.95
f-AnoGAN	57.97
DFR	29.67
DACL	51.23
MFRAD	25.49

表1和表2的结果表明，我们提出的方法在单晶光伏电池数据集上达到了最高的ROC-AUC（0.956）、PR-AUC（0.965）、准确率（0.899）、召回率（0.917）和F1分数（0.902），在多晶光伏电池数据集上达到了最高的ROC-AUC（0.868）、PR-

AUC (0.899)、准确率 (0.810)、和 F1 分数 (0.788)，优于其他模型。原因在于以下几个方面，一方面，与专注于像素级重建的 GANomaly、f-AnoGAN 和 DACL 相比，我们提出的基于多尺度深度特征的方法虽推理速度较慢，但因为能利用丰富的局部细节和多种空间上下文信息在特征空间进行异常检测，极大地提升了检测性能。另一方面，与同样使用深度特征重建进行异常检测的 DFR 相比，我们的方法 MFRAD 利用双重编码结构和基于 Wasserstein 距离的判别器以提高多尺度特征的重构质量，并丰富了异常评分的组成部分，使得其性能比 DFR 更好。

我们的方案在保持高精确率的同时实现了高召回率，能够提供更可靠、实用的检测结果。这是因为该方案有效避免了过多的假阳性和假阴性，这也解释了为什么它在 F1 分数方面优于其他模型。

需要明确指出的是，实际上，现实世界应用的主要关注点是提高检出率，即真阳性率，同时尽可能降低假阳性率 (FPR，通常也称为误报率)。FPR 定义见公式 (18)：

$$FPR = \frac{FP}{FP + TN} \quad (18)$$

因此，我们将所有方案的最大假阳性率设定为 20%，并展示了所有方案的真阳性率随假阳性率的变化情况，如图 4 所示。

其中，上方为单晶光伏电池数据集的结果，下方为多晶光伏电池数据集的结果。

从图 4 可以看出，在相同的误报率，即假阳性率下，MFRAD 具有最高的检出率，明显优于其他三种无监督异常检测方案。

为了直观说明我们提出的方案能够有效地检测有异常的光伏电池，我们提供了测试集上异常分数分布的可视化结果。

其中，上方的是单晶数据集的可视化结果，下方的是多晶数据集的可视化结果。

如图 5 所示，横轴表示归一化的异常分数，纵轴表示在异常分数的每 0.01 区间内正常和异常样本的数量。可以看出，在测试过程中，大多数正常样本的异常得分都很低。相反，异常样本则表现出很高的异常分数。两个分布之间相对较小的重叠也证明了我们模型的优越性和有效性，证实了其能够准确检测出光伏电池中的异常情况。

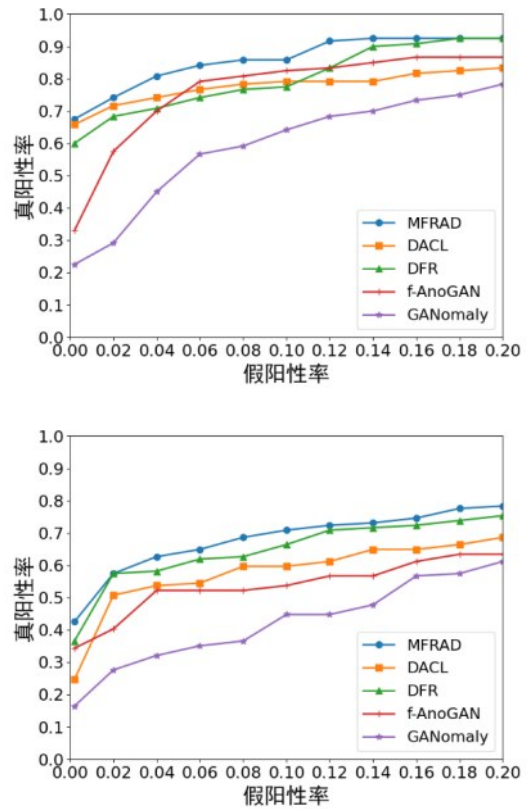


图 4 对比试验的真阳性率 vs. 假阳性率曲线

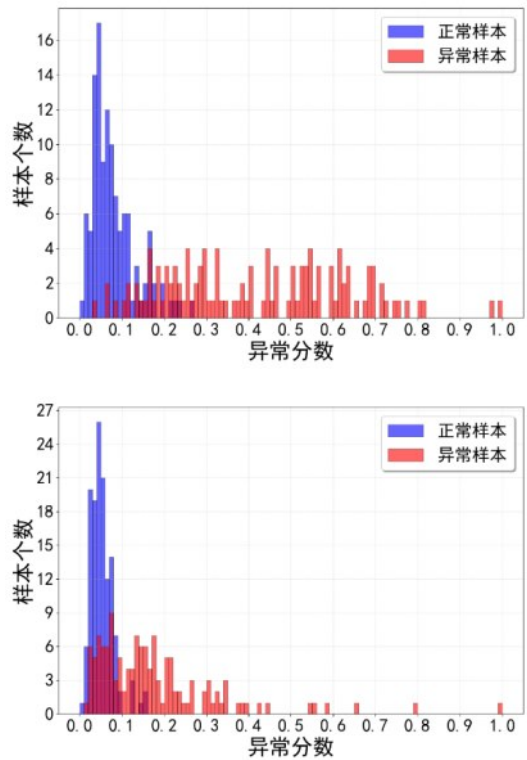


图 5 异常分数分布的可视化结果

为探究实际产线中的域偏移问题，如分辨率的影响，除 256×256 分辨率的输入外，本文还在 128×128 分辨率下进行了实验，实验结果如表 4 所示。可以观察到，128×128 分辨率下在单晶和多晶光伏电池数据集上的检测性能均有所下降。这主要是由于在低分辨率条件下，细粒度缺陷特征被模糊化，导致特征表达的判别能力下降，使得基于特征的异常检测更加困难。

3.4 消融实验

在本节中，我们将进行消融研究，以验证附加编码器和判别器的有效性。我们设计的消融研究如表 5 所示，“-E”表示不使用附加编码器，因此潜空间中的一致性约束。“-D”表示不使用判别器，即不使用对抗策略训练模型。

表 5 MFRAD 变体的结构

变体	组件	附加编码器	判别器
MFRAD-E		×	√
MFRAD-D		√	×
MFRAD		√	√

表 4 单晶和多晶光伏电池数据集在不同分辨率下的实验结果

类型&分辨率	ROC-AUC	PR-AUC	准确率	精确率	召回率	F1-分数
单晶&256	0.956	0.965	0.899	0.887	0.917	0.902
单晶&128	0.921	0.935	0.848	0.839	0.900	0.868
多晶&256	0.868	0.899	0.810	0.888	0.709	0.788
多晶&128	0.798	0.856	0.772	0.848	0.664	0.745

表 6 单晶光伏电池数据集上消融实验结果

模型	ROC-AUC	PR-AUC	准确率	精确率	召回率	F1-分数
MFRAD-E	0.948	0.957	0.891	0.885	0.900	0.893
MFRAD-D	0.943	0.953	0.874	0.888	0.858	0.873
MFRAD	0.956	0.965	0.899	0.887	0.917	0.902

表 7 多晶光伏电池数据集上消融实验结果

模型	ROC-AUC	PR-AUC	准确率	精确率	召回率	F1-分数
MFRAD-E	0.853	0.891	0.809	0.895	0.702	0.787
MFRAD-D	0.849	0.887	0.802	0.862	0.721	0.785
MFRAD	0.868	0.899	0.810	0.888	0.709	0.788

表 6 和表 7 列出了 MFRAD 和其他变体在单晶和多晶光伏电池数据集上的实验结果，最佳性能值

将以粗体标出。可以看出，我们提出的方法在单晶光伏电池数据集上有最高的 ROC-AUC (0.956)、PR-AUC (0.965)、准确率 (0.899)、召回率 (0.917) 和 F1 分数 (0.902)，在多晶光伏电池数据集上有最高的 ROC-AUC (0.868)、PR-AUC (0.899)、召回率 (0.709) 和 F1 分数 (0.788)，总体上优于其他变体。

与上一小节相同，我们将所有方案的最大假阳性率/误报率都设定为 10%，以进一步检验每种方案在实际应用中的性能。从图 6 可以看出，在相同的误报率下，我们的最终方案 MFRAD 具有最高的检出率，优于它的其他变体。

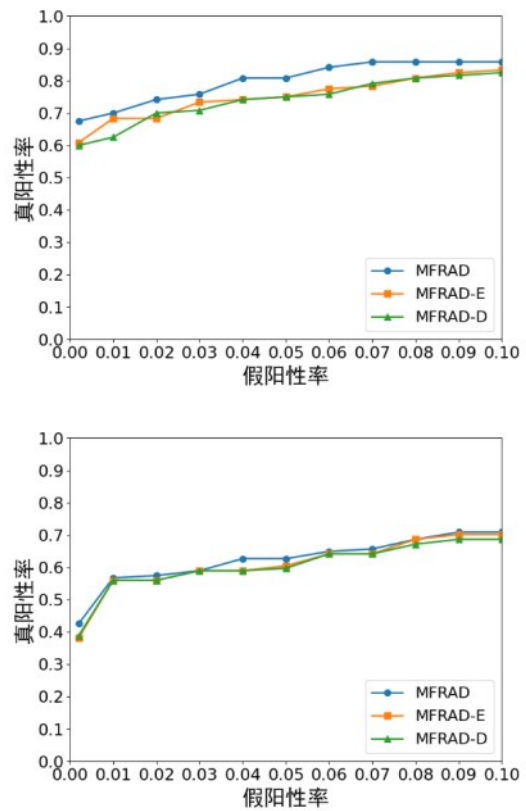


图 6 消融试验的真阳性率 vs. 假阳性率曲线

其中，上方是单晶光伏电池数据集的结果，下方是多晶光伏电池数据集的结果。

可以看出，附加编码器和判别器两个组件对于模型异常检测性能的提升都有作用，且判别器的贡献更为显著。

4 结束语

本文提出了一种基于电致发光 (EL) 图像的

无监督光伏电池异常检测方案。该框架从光伏电池的 EL 图像中提取多尺度深度特征，并在特征空间中利用对抗重构的方法进行异常检测。我们还针对高维的多尺度特征特别设计了重构模块，该模块采用了编码器—解码器—编码器结构的生成器和基于 Wasserstein 距离的对抗训练策略。消融实验显示，双重编码引入的潜在空间一致性约束和对抗性策略均对模型性能的提升有作用。在推理阶段，每个 EL 图像异常评分综合了多尺度特征的重构误差和潜在表示的重构误差，也更好地反映图像的异常情况。相较于 GANomaly、f-AnoGAN 和 DFR，所提方案在单晶和多晶数据集上的 ROC-AUC 分别平均提升了约 3.8% 和 7.0%。同时，各类实验结果进一步验证了所提方案在异常检测任务中的有效性与鲁棒性。

我们的工作可以从以下几个方面进行改进。首先，我们的工作遵循无监督学习范式。虽然无监督方法可以避免标注异常样本的高昂成本，但难以有效利用异常样本的先验知识。因此，值得研究如何利用现有少量异常样本或人工生成的伪异常样本，以实现更有效的异常检测。其次，我们的工作仅针对光伏电池进行异常检测，即只判断光伏电池是否异常，尚未实现异常定位和分割。未来有必要在现有工作的基础上，实现光伏电池异常的定位与分割。

参考文献:

- [1] International Energy Agency. Renewables 2024[R]. Paris: IEA, 2024. [2024-07-29]. Available: <https://www.iea.org>.
- [2] AKRAM M W, LI G, JIN Y, CHEN X, EL-BANBY G M. Failures of photovoltaic modules and their detection: A review[J]. Applied Energy, 2022, 313: 118822.
- [3] AL-OTUM H M. Deep learning-based automated defect classification in electroluminescence images of solar panels[J]. Advanced Engineering Informatics, 2023, 58: 102147.
- [4] OZTURK E, OGLIARI E, SAKWA M, DOLARA A, BLASUTTIGH N, PAVAN A M. Photovoltaic modules fault detection, power output, and parameter estimation: A deep learning approach based on electroluminescence images[J]. Energy Conversion and Management, 2024, 319: 118866.
- [5] Demirci M Y, Bešli N, Gümüşçü A. Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in electroluminescence images[J]. Expert Systems with Applications, 2021, 175: 114810.
- [6] ZHANG J, YANG W, CHEN Y, DING M, HUANG H, WANG B, et al. Fast object detection of anomaly photovoltaic (PV) cells using deep neural networks[J]. Applied Energy, 2024, 372: 123759.
- [7] RAMADAN E A, MOAWAD N M, ABOUZALM B A, SAKR A A, ABOUZAIID W F, EL-BANBY G M. An innovative transformer neural network for fault detection and classification for photovoltaic modules[J]. Energy Conversion and Management, 2024, 118: 118718.
- [8] NGUYEN H, NGUYEN T A, TOAN N D. Optimizing feature extraction and fusion for high-resolution defect detection in solar cells[J]. Intelligent Systems with Applications, 2024, 24: 200443.
- [9] DEMIRCI M Y, BEŞLI N, GÜMÜŞÇÜ A. An improved hybrid solar cell defect detection approach using Generative Adversarial Networks and weighted classification[J]. Expert Systems with Applications, 2024, 252: 124230.
- [10] XIA X, PAN X, LI N, et al. GAN-based anomaly detection: A review[J]. Neurocomputing, 2022, 493: 497 - 535.
- [11] OTAMENDI U, MARTINEZ I, OLAIZOLA I G, QUARTULLI M. A scalable framework for annotating photovoltaic cell defects in electroluminescence images[J]. IEEE Transactions on Industrial Informatics, 2023, 19(9): 9361 - 9369.
- [12] QIAN X, LI J, CAO J, WU Y, WANG W. Micro-cracks detection of solar cells surface via combining short-term and long-term deep features[J]. Neural Networks, 2020, 127: 132 - 140.
- [13] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks[J]. Medical Image Analysis, 2019, 54: 30 - 44.
- [14] GULRAJANI I, AHMED F, ARJOVSKY M, DUMOULIN V, COURVILLE A. Improved training of Wasserstein GANs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS). 2017: 5767 - 5777.
- [15] ZHU J, DENG F, ZHAO J, CHEN J. Adaptive aggregation-distillation autoencoder for unsupervised anomaly detection[J]. Pattern Recognition, 2022, 131: 108897.
- [16] HU D, ZHANG C, YANG T, FANG Q. A deep autoencoder with structured latent space for process monitoring and anomaly detection in coal-fired power units[J]. Reliability Engineering & System Safety, 2025, 261: 111060.
- [17] HADSELL R, CHOPRA S, LECUN Y. Dimensionality reduction by learning an invariant mapping[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2006: 1735 - 1742.
- [18] WANG Z, BOVIK A C, SHEIKH H R, SIMONCELLI E P. Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 - 612.
- [19] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015: 234 - 241.
- [20] KANG X, LI Y, ZHANG Y, MA N, WEN L. Anomaly detection in concrete dam using memory-augmented autoencoder and genera-

- tive adversarial network[J]. Automation in Construction, 2024, 168 (A): 105794.
- [21] AKCAY S, ABARGHOUEI A A, BRECKON T P. GANomaly: Semi-supervised anomaly detection via adversarial training[C]// Proceedings of the Asian Conference on Computer Vision (ACCV). 2018: 622 - 637.
- [22] DENG H, LI X. Anomaly detection via reverse distillation from one-class embedding[C]// Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 9727 - 9736.
- [23] TIEN T D, DO H V, HUYNH T V, KIM D S. Revisiting reverse distillation for anomaly detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023: 24511 - 24520.
- [24] WANG R, LI Z, ZHENG L, WANG W, LI S. Deep feature clustering for multi-class industrial image anomaly detection[J]. Knowledge-Based Systems, 2025, 311: 113134.
- [25] GUDOVSKIY D, ISHIZAKA S, KOZUKA K. CFLOW-AD: Real-time unsupervised anomaly detection with localization via conditional normalizing flows[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2022: 1819 - 1828.
- [26] DEFARD T, SETKOV A, LOESCH A, AUDIGIER R. PaDiM: A patch distribution modeling framework for anomaly detection and localization[C]// ICPR International Workshops and Challenges. 2021: 35 - 42.
- [27] SHI Y, YANG J, QI Z. Unsupervised anomaly segmentation via deep feature reconstruction[J]. Neurocomputing, 2021, 424: 9 - 22.
- [28] TAO D, ZHANG W, MA Z, HOU Z, LU Z, ADAK C. Unsupervised anomaly detection for surface defects with dual-siamese network[J]. IEEE Transactions on Industrial Informatics, 2022, 18 (11): 7707 - 7717.
- [29] SALEHI M, SADJADI N, BASELIZADEH S, ROHBAN M H, RABIEE H R. Multiresolution knowledge distillation for anomaly detection[C]// Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021: 14897 - 14907.
- [30] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2024-07-29]. <https://arxiv.org/abs/1409.1556>.
- [31] LI X, ZHENG Y, CHEN B, ZHENG E. Dual attention-based industrial surface defect detection with consistency loss[J]. Sensors, 2022, 22(14): 5141.

[作者简介]



田家源 (2001-), 男, 南京邮电大学硕士研究生, 主要研究方向为深度学习, 无监督异常检测。



王玉峰 (1974-), 男, 博士, 南京邮电大学通信与信息工程学院教授, 博士生导师, 主要研究数据科学以及人工智能 (数据挖掘、深度学习和强化学习等) 与算法机制设计研究和应用。