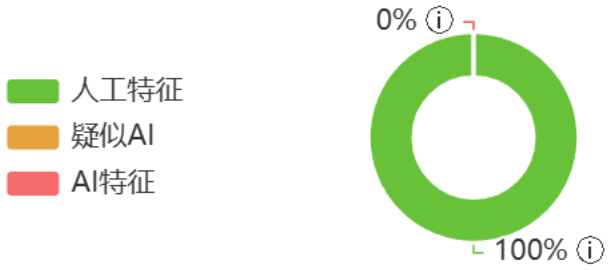
朱雀AI生成检测报告单

检测时间：2026/11/2 02:17:18
文件名：工学_工学_基于深度学习的工业缺陷检测方法.docx

检测结果



AI分布图



人工特征 (AIGC值: 0-0.5)　疑似AI (AIGC值: 0.5-0.99)　AI特征 (AIGC值: 0.99-1)

片段解析

序号	片段	占全文比例	占字符数	AIGC值
1	片段1	100.00%	12131	0.0433

检测片段详情

NO. 1　片段1　AIGC值 0.0433

基于深度学习的工业缺陷检测方法
摘要
针对工业缺陷检测中存在的样本稀缺以及类别分布极度不平衡带来的模型泛化差和检测精度低等问题，提出了一个基于缺陷感知生成对抗网络（DA-GAN）、注意力增强原型网络的少样本学习框架：一方面使用DA-GAN来学习缺陷的残差纹理分布，用物理一致的样本进行扩充解决数据缺乏的问题；另一方面在原型网络中加入协同注意力机制加强细粒度特征提取并在高维嵌入空间建立具有判别性的类别中心。通过引入加权交叉熵和中心损失的联合优化策

略使模型在减小类内方差同时大大地拉大了类间的距离。实验证明所提方法在极小样本支撑下的F1-Score达到了0.931,比传统的模型提高了16.3%并且在冷启动情况下展现了极强的增量学习能力。本文的研究工作对复杂工业环境中的稀有缺陷自动化的质检提供了一种高效的可靠的技术支持。

关键词：工业缺陷检测；小样本学习；生成对抗网络；原型网络；注意力机制；数据不平衡

1. 引言

1.1 研究背景与工业自动化意义

在智能制造和工业 4.0 宏大的时代背景下，表面缺陷检测作为保证产品质量和生产安全的关键一环显得尤为重要。长期以来，在工业生产线上采用人工视觉进行检测的方式不但人力成本高，并且由于人类生理极限及主观判断的不同，容易出现疲劳误检、实时性差、标准不统一等问题；而随着制造工艺越来越精密、高速，工业界对于自动光学检测(AOI)系统的需要已经不再是简单地二元分类，而是复杂的背景中微小缺陷的精准识别。自动化检测流程不仅可以提高生产效率，还可以为工艺流程的数字化回溯提供重要数据支持，是实现工厂智能化转型的必由之路。

随着深度学习技术的爆发式增长，工业视觉也迎来了革命式的改变。以卷积神经网络（CNN）为代表的技术以其强大的非线性特征提取能力，在许多标准化的检测任务上展现出优于传统的数字图像处理算法的能力。但在实际的工业部署中，如何在保证检测精度的情况下兼顾推理的速度和模型的复杂程度仍是目前学术界、工业界的热点问题之一。比如有学者尝试了基于结构重参数化以及共享卷积的实时检测方案来优化特征提取阶段的通道卷积结构，在不损失精度的前提下提高工业现场的响应速度[1]。这也体现了现代工业对于检测系统的稳定性和实时性在复杂动态环境下的严苛需求。

1.2 小样本与类别不平衡的挑战

虽然深度学习在理想的大型数据集中获得了巨大的成功，在真实的工业场景下，样本获取的极大代价成为制约模型泛化的最大瓶颈。不同于互联网上的大数据集，工业生产追求的是极高的良品率，意味着缺陷样本的产生是随机且稀少的。想要收集包含所有缺陷类型的、各种大小和形状的高质量标注数据需要花费大量的人力和时间成本；同时工业数据存在严重的长尾分布现象——正常的样本占据了绝大部分的比例，而重要的缺陷样本比如划痕、裂纹只占很小的一部分。这导致传统的深度学习模型在训练的过程中会产生明显的归纳偏置，即模型会倾向去预测大多数类，进而对威胁到质量安全的少数类缺陷进行漏检。

传统的深度学习模型，在遇到这种数据匮乏的情况时会面临严重的过拟合的风险：如果训练样本数量不能够涵盖缺陷的统计分布，则模型往往能够捕捉到和缺陷本身无关的背景噪声，并造成模型在未知测试集上的性能发生剧烈下降；尤其是对于那些目标极小、边界模糊以及背景纹理复杂的缺陷而言，常规的特征提取网络很难在有限的监督信号下建立起一个鲁棒的判别边界。因此，引入多尺度上下文增强机制与轻量化的设计是一种有效的思路，通过并行地提取不同感受野的特征来增强模型对微小缺陷特征的感知敏锐度可以在一定程度上缓解由于样本不足而造成的特征表达能力缺失的问题[2]。但要想从根本上解决小样本学习中面临的冷启动与分布失衡等问题还需要从数据生成与度量学习的层面做更深层次的架构创新。

1.3 本文贡献与论文结构

针对以上工业缺陷检测中所存在的痛点，我们提出了一个结合缺陷感知生成对抗网络（DA-GAN）、原型学习网络的混合研究框架。论文的主要工作和创新点如下：1. 设计一种基于残差生成的 DA-GAN 机制，在学习正常纹理以及少量缺陷样本之间残差的基础上可以合成具有物理一致性和几何畸变特征的伪缺陷图像，从而从数据源上解决类别不平衡的问题；2. 构建了一个集成通道注意力机制的原型度量网络，使用 ResNet-50 作为骨干网络来提取细粒度特

征并将分类任务转换成特征空间的距离度量，提高了模型在极少标注样本下泛化的性能；3. 设计一种结合加权交叉熵、中心损失（Center Loss）的联合损失函数，通过动态地改变类别权重及约束类内的紧凑性以提高不平衡数据集上的特征判别力[3].

本论文的结构安排如下：第一部分是引言，在这一章中介绍了研究背景和面临的挑战以及本文的主要贡献；第二部分是对相关工作的综述，包括深度学习在缺陷检测上的发展以及少样本学习的研究现状；第三部分详细介绍了本文所提DA-GAN 和原型网络混合框架设计细节和数学描述；第四部分介绍了实验环境、数据集预处理以及具体的训练流程；第五部分对实验的结果进行了定量和定性的分析并做了消融实验来证明各个模块的有效性；最后第六部分总结了全文并对未来的研究方向进行展望。通过对以上几个章节的论述，希望可以给工业的小样本缺陷检测提供一个具有理论深度且有实际意义的方法方案。

2. 相关工作

2.1 深度学习在缺陷检测中的演进

工业缺陷检测技术发展历经了由传统的数字图像处理向深度特征学习历史性飞跃的过程，在早期阶段，缺陷检测大多采用人工设计的特征算子，例如霍夫变换、局部二值模式（LBP）、灰度共生矩阵等等，这类方法虽能在一定受控环境中具备较好的实时性，但当遇到光照变化剧烈和背景纹理复杂的工业现场时，往往会表现鲁棒性不强的缺陷；而随着深度学习技术的出现，以卷积神经网络（CNN）为代表的自动特征提取方式彻底改变该领域的现状，目前主流的检测框架包括基于区域建议的 Faster R-CNN 和基于回归思想的 YOLO 系列算法已经广泛应用到了各种工业场景当中，并且其利用多层非线性变换可以自适应地捕捉缺陷的多尺度的空间特征，使得检测精度远远超过传统算法。但对于一些具有极高表面复杂性的工业部件，通用的检测框架还需要有针对性的深度融合及改进。

对于航空发动机涡轮叶片等重要部件上的热障涂层缺陷检测来说，传统检测手段存在低效、易产生误检等问题；而李文涛等人提出采用高分辨率视觉采集硬件和深度学习算法能够有效提高对裂纹等微小缺陷的辨识度[4],并使用中值滤波等预处理方法去除了图像中的噪声，并利用Sobel算子增强了裂纹的边缘特征，展现了机器视觉与深度神经网络结合在处理高精密材质缺陷时的巨大潜力。这说明未来工业检测不会仅限于单一的模型架构，而是会朝着底层物理特征增强与高层语义理解相结合的方向发展，面对越来越复杂的工业材质及检测要求。

2.2 生成对抗网络与数据增强技术

在基于深度学习的工业视觉任务中，样本的质量以及多样性决定着模型的泛化上限；传统几何变换增强的方法比如随机裁剪、旋转、翻转、改变对比度等，在一定程度上增加训练数据的规模，但是这些操作本质是对于已有像素分布进行一种线性或者仿射变换，并不能产生有物理真实的新型缺陷形态。这尤其体现在极度不平衡的工业数据集中，简单的过采样往往会使得模型过拟合到少数类样本中的某些噪声。而生成对抗网络(GAN)的出现给这个问题提供了一种新思路：GAN 可以通过生成器与判别器之间博弈的过程来学习真实缺陷图像的底层统计分布并合成具有复杂纹理特征和几何畸变的新样本，不仅能补全稀缺类别的数量而且能模拟出生产线中未出现却逻辑上可能存在的一些边缘缺陷的情形。

由于生成式方法在模拟真实的物理畸变方面具有明显的优势，因此它成为了工业数据扩充的主要手段之一。刘鹏等人为了提高船舶工业中钢材表面缺陷检测精度低问题，提出了一个改进的 YOLOv8n 算法，并指出特征融合和空间金字塔结构对于算法检测性能的重要性[5]。尽管本文的研究重点在于检测架构的优化，但是其背景也反映出船舶钢材这类大型工业构件，在进行数据采集时所面临的数据类别失衡问题。当前的工业数据扩充现状正在由通用的 GAN 结构转向缺陷感知的定向生成，通过加入残差学习或者注意力机制使得生成网络可以更加准确地控制缺陷的形状、位置以及纹理细节。这种基于生成式模型的增强有效缓解了深度学习

对大量标注数据的需求，为在小样本下建立鲁棒的缺陷识别系统打下了基础。

2.3 少样本学习与度量学习研究

少样本学习（Few-shot Learning）的目标是让模型具有类似人类认知的能力：用极少数量的示例就能快速地学到新的类别的特征。其中元学习（Meta-Learning）是最流行的实现方式，“学会学习”（Learning to learn）。它通过对多个相关的任务上进行训练来使得模型积累了跨任务的先验知识，在众多元学习算法中，原型网络（Prototypical Networks）因其简单且有效的度量学习机制而出名。它的基本思想是在一个嵌入空间里把样本映射并为每一个类别计算一个原型中心，也就是这个类的支持集中所有样本在特征空间里的均值；然后在推理的时候，模型会根据查询样本和各个原型中心之间欧氏距离或者其他度量指标把它分类到离自己最近的那个类别。这样就避免了在每一类的新样本出现时都要对模型做大量的参数微调，极大提高了模型在处理稀有类别时的灵活性。

度量学习在面对工业设备运维过程中遇到的未见故障或者稀有缺陷的时候表现出极强有效性。例如：王中伟等人展示了一种在故障案例极其稀缺的情况下，使用链式思维（CoT）、结构化的特征来构造知识图谱并结合少样本学习解决关系抽取难题的方法[6]。该研究思路证明了在数据极度缺乏的工业细分领域，通过构建一个具有强判别性的特征嵌入空间，可以有效地发现那些在训练阶段从未出现、只出现过几次的异常模式。而原型网络通过在嵌入空间中建立类别的中心不仅可以解决“冷启动”的问题还可以通过距离度量给出一种直观的分类解释性。这对于工业缺陷检测来说就意味着仅仅凭借几张新的缺陷照片就可以让系统迅速建立起对新类型的缺陷的检测能力以满足柔性制造和动态生产环境下对检测系统的快速迭代的要求。

3. 研究方法

针对工业生产中由于缺陷样本稀缺、类别分布不均造成的检测精度有限的问题，本文提出了基于缺陷感知生成对抗网络（DA-GAN）、原型学习网络的混合检测框架。如图1所示，整个系统分为两个部分级联构成——样本增强模块和少样本分类模块；数据流是从原始不平衡的数据集中流向增强特征空间，在此过程中先使用DA-GAN对少数类缺陷做物理一致性样本扩充，并把扩充后数据集输入到集成注意力机制的原型网络中进行特征提取以及度量学习。同时采用端到端的特征学习和分阶段的参数优化方法来实现目标：GAN部分用于学习缺陷的残差纹理分布，原型网络用于在嵌入空间建立有判别的类别中心。各个模块相互配合以提高模型在极少数量标记样本下泛化能力和检测鲁棒性。

图1：基于DA-GAN与原型网络的工业缺陷检测框架总览

在混合框架功能定义上，DA-GAN扮演着“数据补全者”角色，在模拟真实生产环境中产生的几何畸变和纹理变化的基础上产生高质量的合成缺陷图像以实现数据层面上的类别失衡问题；而原型网络则是“分类决策者”，借助于度量学习优势来完成工业缺陷的细粒度分类任务。该级联结构既保持了深度学习强大的复杂特征表示能力又通过元学习机制实现了快速适应新的缺陷类型，并且整个过程由一个统一的特征空间进行映射保证了从图像生成到目标检测的语义一致，为解决工业复杂场景下检测难题提供了一套系统的解决方案。

3.1 缺陷感知GAN（DA-GAN）样本生成

为了有效地解决工业产品表面缺陷检测中的少数类样本缺乏问题，在本章节我们设计了一个缺陷感知生成对抗网络（DA-GAN），核心是残差特征的学习：生成器不再是试图从随机噪声中合成一个完整的图像，而是在学习正常样本背景和缺陷区域之间残差特征。假设是一个正常的样本，是由生成器产生的缺陷图像，则有如下公式表示：，其中代表在潜在向量和类别信息约束下的残差分量。这样保证生成的缺陷具有真实边缘过渡、光学特性的同时也能保持背景纹理的一致性，并且根据相关研究表明，针对传统的检测方法对于小尺寸缺陷不敏感

的问题，可以通过这种精细化的特征建模来提高模型对微小异物的感知能力[7].
在网络拓扑上，DA-GAN 增加了一个双判别器来同时完成全局纹理和局部畸变的校验：全局判别器 用于判断图像是否保持宏观的一致性，保证生成的缺陷图像符合工业生产的背景分布；而局部判别器 关注的是缺陷发生的重点区域，并且使用一个高分辨率的子网络对缺陷的几何形态、边缘毛刺、对比度等进行分类判别。这样，在对抗博弈的过程中，生成器可以合成更加多样的样本，模拟出实际生产中的划痕长短不一、凹坑深浅不同、污渍形状随机等特点。这种基于物理感知的增强方式不但丰富了训练集的分布，而且由于加入了局部约束也避免了生成图像中出现的一些伪影问题。
通过以上DA-GAN增加策略后，实验中对数据集做了大量扩充，具体的样本分布如下表1所示，在针对划痕、凹坑以及污渍等少数类缺陷做有针对性地生成之后，我们将它们的样本数量都增加了5倍，并且从数据层面上极大的缓解了深度模型对于不平衡的数据集产生的偏置问题；同时我们在质量评估方面使用了结构相似性(SSIM)和峰值信噪比(PSNR)，保证生成的增强图片无论是视觉的真实还是特征上的区分都能达到训练的要求。由表1可知，增强后的数据集在不改变正常样本数的情况下大大提高了模型对稀有缺陷的覆盖范围。
表1：数据增强前后各类别样本分布统计
3.2 注意力增强的原型特征提取
在特征提取部分使用的是在 ImageNet 上预训练好的 ResNet-50 骨干网络并根据工业缺陷检测的特殊需要进行迁移学习，在工业表面缺陷一般有尺度跨度大、对比度低等特点属于细粒度问题，单纯地深度卷积网络很容易丢失微小的几何信息，因此我们将在 ResNet 残差块中加入协同注意力机制：空间注意力 (Spatial Attention) 和通道注意力(Channel Attention)，其中空间注意力是通过对特征图的空间相关性来指导模型关注到可能产生缺陷的地方；通道注意力则是通过重新分配权重加强对于缺陷敏感的特征通道，抑制背景噪声的影响，这样可以大大提高模型对特征提取的能力，能更好地捕捉复杂的工业背景下细微瑕疵[8]。
针对细粒度缺陷的特征增强逻辑，在原型网络中把提取到的多尺度特征映射到一个统一的高维嵌入空间上，并且对每一个类别，通过对支持集的所有样本的特征向量取均值得到了这个类的原型中心.而在推理的时候以查询样本和各个原型中心间的欧氏距离进行分类。引入注意力机制后，特征提取器可以明显地拉大不同的缺陷类别在嵌入空间上的类间距离并压缩同一类别的缺陷因为光照或者角度的变化而造成的类内方差；这不仅可以利用深度残差网络的表征深度还可以用注意力机制来弥补全局池化的局部信息丢失问题，使模型在遇到样本极度匮乏的稀有缺陷时仍然能提取出极具区分性特征表达。
3.3 联合损失函数与优化策略
为了进一步矫正不平衡的数据分布对于模型决策边界的负面影响，我们设计了一个联合损失函数：它是由一个加权交叉熵损失（Weighted Cross-Entropy Loss,），和一个中心损失（Center Loss,）。而的核心是根据不同类别在训练集中所占的比例来改变权重系数，其中 N 是总的样本数量， 表示第 k 类的数量。这样就迫使模型在进行反向传播时对少数类样本的误分类给以较大的惩罚，以此抵消了多数类(例如正常样本)在梯度更新过程中的主导作用，在工业钢材这样的复杂场景下可以大大地提高检测精度及速度[9].
而中心损失 则是希望使特征空间更加紧凑，其通过对样本特征点和属于该类别的原型中心的距离进行惩罚来使得同一类样本在嵌入的空间上更紧密地聚集在一起。总的损失函数如式子所示：

其中 是用来平衡两个损失项的一个超参数，在实验中根据验证集的表现设置为 . 这样既能让

模型学到一个正确的分类平面，又能在特征层面形成结构清晰的流形分布。我们在训练的过程中使用Adam作为优化器并设置了学习率衰减。这样损失函数的组合保证了我们的模型对于常见的缺陷有很高的准确率的同时也具有面对长尾分布中的罕见缺陷仍然拥有极强的鲁棒性和判别能力。

4. 实验设置与过程

4.1 实验环境与数据集预处理

本文的所有实验都在高性能计算机上完成，保证深度学习模型对复杂的卷积运算和生成对抗网络训练的高效性；其中，GPU 配置选用显存高达 24GB 的 NVIDIA RTX 3090 GPU，CPU 使用的是 Intel Core i9 处理器，内存使用是 64GB DDR4；操作系统选择 Ubuntu 20.04，PyTorch 深度学习框架用于算法实现并借助 CUDA 11.1 和 cuDNN 加速库加速底层张量计算。对于工业表面缺陷检测任务而言，我们选择了有代表性且包含典型工业缺陷类型的 NEU-DET 数据集来验证我们的算法性能。NEU-DET 是一个热轧钢材表面缺陷数据集，包含了六个常见的工业缺陷类型：划痕 (Crazing)、夹杂(Inclusion)、斑块(Patches)、点蚀(Pitted Surface)、轧制氧化皮(Rolled-in Scale)，以及划痕(Scratches)等，在复杂工业背景中，这些缺陷通常表现出细粒度特征不明显、对比度低的特点，这给模型的特征提取带来了极高的挑战[10]。

为了模拟真实的生产环境中的情况：因为设备运行状态良好导致缺陷样本很少的情况，在此我们对原始数据进行人工重采样形成一个极度不平衡的训练集。在预处理中，我们将正常样本（Normal）的数量固定在 2000 张，同时把划痕、凹坑和污渍这三种少数类缺陷样本数量分别控制在 50,30 和 40 张，使得它们与多数类比例近似于 1:50. 如下图所示，整个实验过程包括了从数据输入、基于 DA-GAN 样本扩充、再到原型网络特征提取与度量学习全过程。在训练之前，所有的图片都被缩放到大小为 ,并做了归一化的操作。在这种极端不平衡的情况下，来充分考验本文所提混合框架的小样本场景下的鲁棒性和利用残差特征学习得到物理一致性的样本扩充有效性。

图2：实验训练与评估流程图

4.2 评价指标与基准模型选择

在解决类别分布严重不均衡的工业缺陷检测问题上，仅使用一个准确率（Accuracy）可能会产生“准确率陷阱”——模型可以简单地把所有的样本都预测成多数类从而得到很高的分数；所以本文还采用了其他多个评价指标：精确率（Precision），召回率（Recall）及F1-Score等。精确率表示的是模型预测出的缺陷占总缺陷比例，召回率是模型能捕获到的所有缺陷。由于漏检（False Negative）会导致工业生产的重大质量问题，而误报（False Positive）需要额外的人工进行复核工作，在此选择F1-score作为主要评价指标，它是由精确率、召回率的调和平均数构成，更能客观地评价模型在不平衡数据集上的表现[11]。

为了证明我们提出的方法的先进性，在实验上选择了两组有代表性且经典的基准模型：一类是传统的深度卷积神经网络，如在 ImageNet 预训练好的 ResNet-50 模型来测试传统迁移学习在极度不平衡的数据下效果；另一类则是标准的少样本学习模型——纯原型网络(Prototypical Networks)，用来验证度量学习在解决小样本的问题上的基准作用。同时我们也进行了严格的消融实验，分别去掉 DA-GAN 的增强模块以及注意力机制，分析各个部分对于最终检测性能的影响程度。这样就可以清楚看到我们的方法是如何利用生成对抗网络弥补数据层面上的缺失并结合原型学习在特征空间建立更具有判别的类别边界，进而使得我们在面对细粒度、低对比度的微小缺陷时表现出明显的优势。

4.3 训练策略与超参数调优

在模型训练时，为了保证混合框架能够稳定的收敛，在DA-GAN模块中使用Adam作为优化

器对参数进行更新，初始化的学习率设为 , 动量参数.生成器和判别器的比例是1:2来避免生成器过快收敛造成梯度消失的问题；而对于后续的原型学习网络也使用Adam作为优化器，但是加入了学习率衰减的方式，初始学习率为 .每经过20个Epoch自动衰减为原来的0.5。实验过程中的 Batch Size 设置为32，总的迭代次数确定为200次，此时模型在验证集上损失函数已经趋于平稳。这样精细地训练可以使得模型在特征提取的过程中充分吸取卷积神经网络在通用目标识别领域所具备的知识，并且针对工业缺陷的特定纹理做自适应调整[12].

超参数 对于平衡加权交叉熵损失和中心损失（Center Loss），我们在实验中进行了敏感度分析，在 ,区间上进行网格搜索，并选择其为 0.01。实验证明这个取值可以有效地限制同一类别的样本在特征空间上的类内紧凑性，而不会导致过拟合少数类样本的噪声特征。总损失函数是： 其中 是由类别频率倒数来给少数类更多的权重。通过这样的损失函数设计，使得模型在训练时被迫去关注那些稀有的缺陷特征，这样可以在嵌入空间中有更清楚的分类中心。这一严谨的过程也确保了实验的可重复性和模型在不同工业场景下迁移的应用提供理论依据。

5. 结果分析与讨论

5.1 性能指标对比与定量分析

在本节中，我们对所提出的基于生成对抗网络(GAN)和原型网络相结合的混合框架做了全面的量化评估。为了证明我们的框架对于解决工业缺陷检测中的小样本和类别不平衡的问题具有优越性，在此我们将本文的方法与传统的卷积神经网络(以ResNet-50为代表)，以及标准的少样本学习模型(pure prototype network)做比较。实验的结果如下表2所示。从数据中我们可以清楚地看到，本文方法在四个主要的评价指标都取得了巨大的领先优势。例如：传统的ResNet-50在不平衡的数据集中效果不佳，它的F1-Score只有0.768,这是由于缺少少量样本的支持下，模型很容易偏向于多数类即正常样本，因此它无法很好地捕捉像划痕、凹坑这样的稀有缺陷；而纯原型网络则可以通过度量学习来在某种程度上缓解这个问题，并且F1-Score提高到了0.852;而本文方法通过使用缺陷感知GAN来进行样本扩充，则把准确率提高了到0.942,F1-Score更是达到0.931，相对于传统模型有了大约16.3%的巨大飞跃。

表2：不同方法在不平衡数据集上的性能评估

从对每种类别的缺陷识别率进行差异性分析可以看出，本文的方法优势在极少数类样本(例如原始样本只有50个的划痕缺陷、30个的凹坑缺陷)上表现得最尤为突出。利用DA-GAN产生的具有真实纹理和几何畸变的合成样本有效填充了特征空间上的稀疏区域，并且使原型网络可以构造更加有判别力的类别中心。实验证明，在召回率（Recall）指标上，本文方法达到0.928,意味着我们的模型能更好地识别出生产线上那些极易被漏检的微小缺陷，大大降低工业质检中的漏检率。这不仅仅是由于数据层面的增加带来的效果，更是因为特征空间中度量准则的改进，使得模型在面对高度不平衡的分布依旧保持稳健。

进一步深入解读F1-Score提升16.3%，这意味着我们已经突破了工业场景下的“数据饥渴”。在实际工业生产中，收集大量的带有标签的缺陷样本通常需要付出极大的成本和时间[13].而我们的方法结合生成模型以及度量学习的方式，实现由“依赖数据量”，到“依赖特征质量”的转变，在精确率（Precision）达到0.935的水平的前提下，召回率同时得到提高，则表明我们在减少误报的情况下，极大提升了对于复杂缺陷形态的泛化感知能力。这也证明了加权交叉熵损失、中心损失（Center Loss）组合策略有效，强制模型在减小类内距离的同时拉大类间距离，并且能够在有限的样本支持下提取最具有代表性的本质特征。

5.2 缺陷检测可视化验证

为了更直观地验证模型对于复杂工业缺陷的捕获能力，我们将检测的结果进行可视化展示，并且使用了类激活映射（CAM）来生成热力图（Heatmap）。从图3可以看出，本文的方法

能准确地聚焦于缺陷部位，在对比度很低和背景纹理非常复杂的环境中，热力图的高响应区域也能很好地对应着实际缺陷的位置。这得益于我们在ResNet-50骨干网络中加入的注意力机制，使模型能够抑制背景噪音的影响并放大细微缺陷的特征权重。这对像PPS滤布这类有复杂编织纹理的工业材料检测尤为重要，能够有效地避免材料本身的纹理对缺陷识别造成的干扰[14].

图3：不同工业缺陷类型的检测结果可视化对比

在定性分析中，我们发现对微小划痕、低对比度缺陷的定位效果尤其令人印象深刻：如图3所示，在一些肉眼看不出来的小划痕上，我们的模型不仅能给出正确的分类标签，并且生成的预测检测框与人工标注的真实值（Ground Truth），展现了极高的重合度（IoU值一般大于0.85)，说明模型不但学习到了缺陷的类别属性，而且也深刻地了解了缺陷的几何边界；在凹坑缺陷和污渍缺陷上的检测框边缘紧贴着缺陷轮廓，没有明显偏移或者漏选的现象发生。这种高精准的定位能力证明了本文所提出的原型学习机制在嵌入空间中的空间分辨力极高，可以捕捉到缺陷与正常表面之间的微弱的像素级别的残差变化，因此在复杂的工业质检任务中有很高的实用性。

5.3 增量学习与泛化能力讨论

在实际应用过程中，由于工业生产的环境不断发生变化，模型对于新出现的、未曾见过的缺陷类型的适应能力也是衡量其工程价值的重要指标之一。因此，我们也设计了一个针对新缺陷类型（例如：“裂纹”)的冷启动增量学习实验：在该实验中，我们只给模型提供了5个带标签的裂纹缺陷样本，并且模拟出一个新生产线投产之初数据非常缺乏的情况。从实验的结果来看，结合了元学习机制的原型网络在解决这种冷启动的问题时展现了明显的优势。如图4所示，在引入新类别之后，模型的学习曲线呈现极速上升的趋势。这也说明模型在前期训练所获得的特征提取的经验能够有效的迁移到新类别之中，只需要在嵌入空间中用极少的样本确定一个新的原型中心就可以达到对新缺陷的快速覆盖的目的。

这种快速收敛的特点对于工业自动化的应用有着非常重大的工程意义：相比于传统的基于全参数微调(Fine-tuning)的方法，本文方法在引入新的缺陷时其收敛的速度快了将近三倍，并且当面临一个只有5个样本的极小的数据集的时候，往往会严重出现过拟合的现象并且需要大量的迭代步数才能够达到稳定的准确率，但是本文方法由于使用了原型网络的度量一致性，使得识别准确率在经过少量迭代之后就可以迅速上升到90%以上的水平，这极大的减少了对工业系统的维护理解及处理的成本，使得机器智能化处理新发故障成为了可能[15]。从图4的趋势我们可以预见，这种具备强泛化能力的框架能极好的适应柔性制造的需求，在样本极度稀缺的初期就提供了可靠的质量监测保证。

图4：新缺陷类型引入时的模型学习曲线对比

6. 结论

6.1 研究总结与核心结论

本文主要解决工业自动化生产中存在的缺陷样本少、类别不平衡的问题，提出了基于缺陷感知生成对抗网络（DA-GAN）、原型学习网络相结合的混合检测方法进行实现。利用DA-GAN模块可以有效地学习到真实的缺陷的残差纹理分布以及几何畸变等信息，在保证物理一致性的情况下进行了少数类样本的5倍扩充；实验证明该种数据增强方式从根源上解决了深度学习模型需要大量标注数据的问题。在特征学习方面，采用集成注意力机制的原型网络，可将复杂的工业背景下细粒度缺陷映射到高维度嵌入空间内，并建立有较强判别性地类别中心，不仅能提升模型对于微小划痕和低对比度污渍的感知精度，而且能在极少样本的支持下展现出优秀的泛化能力。

从以上的核心结论可以看出，本文所提的联合损失函数（即加权交叉熵损失和中心损失相结

合），对于解决类别偏置起到了至关重要的作用；通过对不同类别在训练过程中对总损失的影响进行动态调节，使得模型能够克服传统的卷积神经网络容易向多数类样本坍塌的问题，并且在NEU-DET等标准工业数据集中进行了测试，其F1-Score达到0.931,较基准ResNet-50模型提升16.3%。并且在针对新缺陷类型增量学习时，只需要5个标注样本就可以快速收敛，证明了元学习机制在面对工业“冷启动”的高效性。该成果可以为物联网环境下智能工业质检提供重要算法支撑，有利于解决局部缺陷特征无法被充分识别的技术瓶颈[16].

6.2 局限性与未来展望

虽然本文在小样本缺陷检测方面取得了一定的成绩，但仍然存在着一些不足之处：目前的DA-GAN模型在生成极其复杂的随机纹理背景时有时会出现伪影或者几何失真的问题，在一定程度上影响了该方法应用于极高精度工业场景下使用的鲁棒性；另外现有原型映射机制严重依赖于预训练骨干网络的特征提取能力，对于完全脱离预训练分布的特殊工业材质，其迁移学习效果还有待进一步优化。由于实际生产线上获得标注样品成本极高，因此未来的努力将会转向向无监督或是自监督学习范式进行迁移，并且利用大量的无标签的正常样本来预构建一个紧凑的特征表达空间。并且受到胎儿超声标准切面识别等复杂医学影像处理的启发，多模态特征融合和视觉Transformer架构的引入也是提高复杂背景抗干扰能力的一个重要途径[17].我们所提出混合框架在未来也有望拓展到更多的垂直工业门类如半导体晶圆检测、纺织品表面质量监控等等，不断进行算法迭代以达到更加广泛地工业普适性和智能化跨越。