

面向领域偏移工业异常检测的分布对齐分数融合与最大池化

翁长明, 杨 艺

江苏理工学院机械工程学院, 江苏 常州

收稿日期: 2026年8月11日; 录用日期: 2026年9月7日; 发布日期: 2026年9月17日

摘 要

领域偏移会改变工业检测系统的正常特征分布及异常分数尺度, 使实验室条件下有效的检测器在部署后产生不稳定判定。针对教师-学生蒸馏分数与原型匹配分数难以直接比较、稀疏缺陷易被全局聚合稀释以及后处理参数耦合等问题, 文章提出一种无需重新训练骨干网络的推理时决策层。该方法基于正常训练集统计量对异构分数进行位置-尺度标准化与加权融合, 采用最大池化保留局部高响应, 并以单一平滑参数控制图像级筛查与像素级定位之间的权衡。在AeBAD-S上, 完整配置取得88.48%的图像级AUROC, 较MMR的84.7%提高3.78个百分点; 当 $\sigma = 1.5$ 时, 像素级AUROC为83.90%, 较MMR低1.40个百分点, PRO为76.03%。结果表明, 该方法更适合以产品级合格/不合格判定为优先目标的部署场景, 其收益伴随区域覆盖能力下降。因此, 文章将其定位为现有工业异常检测流程的轻量化决策层, 而非通用分割模型。

关键词

工业检测, 异常检测, 域偏移, 分数融合, 位置-尺度标准化, 产品级筛查

Distribution-Aligned Score Fusion and Max Pooling for Industrial Anomaly Detection under Domain Shift

Changming Weng, Yi Yang

School of Mechanical Engineering, Jiangsu University of Technology, Changzhou Jiangsu

Received: August 11, 2026; accepted: September 7, 2026; published: September 17, 2026

Abstract

Domain shift changes the normal-feature distribution and anomaly-score scale of industrial inspection

文章引用: 翁长明, 杨艺. 面向领域偏移工业异常检测的分布对齐分数融合与最大池化[J]. 传感器技术与应用, 2026, 14(5): 861-879. DOI: 10.12677/jsta.2026.145083

systems, causing detectors that perform well under laboratory conditions to produce unstable decisions after deployment. To address the limited comparability between teacher-student distillation scores and prototype-matching scores, the dilution of sparse defects by global aggregation, and the coupling of post-processing parameters, this paper proposes an inference-time decision layer that requires no backbone retraining. The method performs location-scale standardization and weighted fusion of heterogeneous scores using statistics estimated from normal training data, applies max pooling to preserve strong local responses, and uses a single smoothing parameter to control the trade-off between image-level screening and pixel-level localization. On AeBAD-S, the complete configuration achieves an image-level AUROC of 88.48%, exceeding MMR's 84.7% by 3.78 percentage points. At $\sigma = 1.5$, the pixel-level AUROC is 83.90%, 1.40 percentage points below MMR, while the PRO score is 76.03%. These results indicate that the method is better suited to deployment scenarios prioritizing product-level pass/fail decisions, with reduced region-coverage performance as the corresponding trade-off. The proposed method is therefore positioned as a lightweight decision layer for existing industrial anomaly-detection pipelines rather than as a general-purpose segmentation model.

Keywords

Industrial Inspection, Anomaly Detection, Domain Shift, Score Fusion, Location-Scale Standardization, Product-Level Screening

Copyright © 2026 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

工业质量控制日益依赖自动化视觉检测系统,以降低人工负担并形成一致的产品级合格/不合格判定。然而,光照、相机位置、背景纹理和产品外观在部署阶段均可发生变化,导致训练域与运行域之间出现偏移。由此,实验室条件下表现良好的检测器仍可能在生产线上产生不稳定的异常分数和报警结果。

现有工业异常检测研究主要从特征表征入手,包括单类学习、教师-学生蒸馏、重建和记忆库匹配等。AeBAD与分布偏移异常检测基准进一步揭示了光照、视角和背景变化对检测性能的影响[1][2]。但在骨干网络输出异常证据之后,分数处理方式同样会影响最终决策:融合模块需要比较数值尺度不同的异常线索,聚合模块则需要将稠密响应转换为图像级分数。若直接对异构分数加权求和,幅值较大的分数源可能主导结果;若采用全局平均池化,面积很小的缺陷响应又可能被大量正常像素稀释。

针对上述问题,本文提出分布对齐分数融合(Distribution-Aligned Score Fusion, DASF),并将其定位为面向异常检测的后处理领域自适应模块。本文所称“分布对齐”具有明确的统计边界:它利用正常训练集统计量,对蒸馏分数和原型分数分别执行基于一阶矩(均值)与二阶中心矩(方差,以标准差表示)的位置-尺度标准化,以降低两类分数的量纲和尺度差异,而不匹配高阶矩,也不等同于完整概率分布对齐。在此基础上,该模块引入正常特征库(Normal Feature Bank, NFB)补充原型证据,通过最大池化保留稀疏异常的峰值响应,并依据高斯卷积性质将后处理简化为单一平滑参数。全部操作均位于固定源模型的特征提取之后,无需访问目标域标签或修改源模型参数,旨在提高检测器面对未知运行域偏移时的产品级泛化能力。本文不将这一结果扩展为适用于所有异常检测架构的通用结论。

围绕上述问题,本文的主要贡献如下:

- 提出一种面向领域偏移工业异常检测的后处理领域自适应模块。该模块将 NFB 原型证据、基于正常训练统计量的一阶矩与二阶中心矩标准化、图像级证据聚合及单参数平滑组织为统一流程；在不访问目标域标签、不更新教师 - 学生源模型的前提下，提高异构异常分数的可比性与模型在运行域中的产品级泛化能力。其输出是经位置 - 尺度标准化的异常分数，而非概率校准后的缺陷置信度。
- 构建与各组件相对应的消融证据链。在固定下游设置下，将 NFB 融合权重由 $\alpha = 0$ 调整为 $\alpha = 0.2$ ，图像级 AUROC 由 86.65% 提高至 87.83%，像素级 AUROC 仅变化 -0.05 个百分点。对于池化策略，本文同时报告控制其余设置的直接比较和逐步配置比较，并将完整流程相对“仅蒸馏 + 平均池化”的 11.28 个百分点增益明确解释为 NFB、池化与平滑的联合效果。
- 给出面向部署的性能边界与运行选择。在 AeBAD-S 上， $\sigma = 3.0$ 时图像级 AUROC 为 88.48%，较 MMR 提高 3.78 个百分点； $\sigma = 1.5$ 时像素级 AUROC 和 PRO 分别为 83.90% 和 76.03%。MVTecAD 上最大池化相对 avg_max 的图像级增益仅为 0.30 个百分点，说明该策略的收益与数据条件和任务目标相关，不能概括为普遍优势。

2. 相关工作

2.1. 工业视觉质量检测

工业视觉质量检测利用传感技术与算法决策支持识别缺陷产品，并辅助完成合格/不合格筛查。工业图像异常检测通常被表述为单类学习问题，即训练阶段仅提供正常样本，而推理阶段同时要求完成图像级检测和像素级定位。MVTec AD 已成为该任务的标准基准，其扩展数据集进一步引入逻辑约束、三维数据以及更复杂的工业场景[3]-[5]。近期综述系统总结了特征学习、重建、蒸馏和记忆库方法在工业检测领域的快速进展[6] [7]。

教师 - 学生蒸馏与原型匹配是两类常见的异常评分机制。前者以冻结教师网络的特征为参照，利用正常样本训练学生网络进行逼近，并将测试时的教师 - 学生差异作为异常证据。Uninformed Students、STFPM、MKD、Reverse Distillation、AAND、HKT 和 EfficientAD 均遵循这一基本思想，并在特征层级、重建路径和计算效率方面作出不同设计[8]-[14]。后者存储或建模正常特征模式，通过测试特征与正常原型之间的距离进行评分，代表性方法包括 PaDiM、PatchCore 和 SPADE [15]-[17]。CFA、DMAD、FastRecon、OmniFrequency 和 GRAD 进一步改进了特征适配、记忆构建或稠密匹配[18]-[22]。这些方法持续提升特征表征能力，但其部署效果还取决于异常分数如何被标准化、融合与聚合。

近年来，研究方向进一步由逐类别模型扩展到统一模型、零样本和少样本检测。UniAD 以单一模型处理多个工业类别[23]；WinCLIP 利用视觉 - 语言预训练模型进行零/少样本异常分类与分割[24]；AnomalyCLIP 通过对象无关提示学习增强跨类别零样本识别[25]；PromptAD 则仅使用正常样本学习少样本异常提示[26]。这些方法侧重提升表示的通用性，本文则在固定骨干网络的条件下研究异构分数的对齐、融合与产品级聚合。

2.2. 领域偏移条件下的异常检测

同域基准无法充分反映真实部署条件。在实际检测系统中，光照变化、传感器位置调整、表面纹理变化和背景偏移均可能改变正常样本及异常分数的分布。Cao 等从一般异常检测角度建立了分布偏移评估框架[1]；AeBAD 则针对航空发动机叶片构建了包含光照和视角变化的工业基准[2]。此外，VisA、IM-IAD、Texture-AD 和 Real-IAD 分别从类别复杂性、制造场景、纹理变化和多视角采集等方面扩展了

真实工业评估条件[27]-[30]。这些工作共同表明, 部署可靠性不仅取决于特征提取器, 还取决于后续分数处理对运行条件变化的敏感性。

当系统联合多种异常线索时, 这一问题更为突出。教师-学生特征差异与正常原型匹配分别刻画模型不一致性和正常模式偏离程度, 二者可能具有不同的经验分布与数值尺度。多数既有研究主要评价特征学习方法本身, 本文则关注这些异构分数在部署阶段的可比性及其对产品级决策的影响。

2.3. 分数融合与校准

在统计学习领域, 分数校准长期以来被研究为将模型原始输出转换为可比较、可解释置信分数的问题。Platt 缩放利用 Sigmoid 模型映射支持向量机输出, 后续研究则采用分箱、保序回归和 Beta 校准等方法改进概率估计[31]-[33]。从统计学习视角看, 校准将分数生成与分数解释相分离; 当不同模型或特征空间产生的数值尺度不兼容时, 这一点尤为重要[34]。在异常检测领域, 研究者也探索了离群分数的归一化与统一, 以增强不同方法和数据集之间异构离群分数的可比性[35]。

数据融合为异构检测证据的组合提供了技术基础。经典多传感器数据融合强调跨传感器或算法的对齐、不确定性处理与决策级组合[36] [37]。分类器组合理论还表明, 求和、乘积、最大值和加权平均等简单规则, 只有在分类器输出具备可比性时才能有效发挥作用[38]。在工业异常检测中, 包括 G²SF 和跨模态特征映射在内的近期多模态、多源方法展示了融合的实际价值[39] [40]。然而, 在教师-学生检测流程中, 如何以低开销方式提高异构异常分数的可比性, 并将其稳定地聚合为产品级判定, 仍缺乏针对性的讨论。

DASF 的研究边界可概括为: 与 MMR [2]相比, 本文保留其骨干网络训练目标和教师-学生重建路径, 仅在特征提取之后增加分数标准化、原型证据融合和图像级聚合。与 G²SF [40]等多模态方法不同, 本文处理的是同一单模态检测流程产生的两类异常分数。与经典分类器组合[38]及 Platt 缩放、保序变换和 Beta 校准[31]-[33]相比, DASF 不进行类别投票, 也不估计缺陷概率; 其 Z-score 变换仅用于降低两个分数源之间的位置与尺度差异。这一处理与离群分数统一研究[35]在提高可比性方面相关, 但本文进一步关注稠密分数融合及其产品级聚合。表 1 总结了上述区别。

Table 1. Positioning relative to closely related work
表 1. 与密切相关工作的定位比较

相关研究方向	主要关注点	本文的区别
MMR [2]	面向工业异常检测的领域偏移数据集与掩码多尺度重建	保留 MMR 式骨干网络, 改进分数对齐、原型证据、池化与平滑
G ² SF [40]	几何引导的多模态分数融合	采用单一检测流, 融合蒸馏与原型异常分数, 而非模态特定的几何线索
分类器融合[38]	分类器输出的通用组合规则	处理分布不匹配的稠密异常分数图与图像级筛查决策, 而非类别标签投票
分数校准[31]-[33] [35]	概率校准或离群分数可比性	仅以 Z-score 对齐降低两个异常分数源之间的尺度支配, 不估计经校准的缺陷概率

2.4. 面向微小缺陷的决策聚合

生成像素级异常图后, 检测系统需要将局部响应聚合为图像级分数。平均池化稳定且易于实现, 但它刻画的是全局平均异常程度。对于仅占少量像素的工业缺陷, 局部高响应会被大量正常区域稀释; 领

域偏移若同时抬高正常背景分数, 还会进一步降低图像级判定的区分度。

最大池化和前 k 项聚合能够更直接地保留局部异常证据。StructCore 研究了结构感知的图像级评分与前 k 比例特征匹配, 极值理论则为利用样本极值刻画尾部事件提供统计依据[41] [42]。与此同时, 高斯平滑、边界细化和阈值化会改变异常图的峰值、边界与覆盖范围, 进而影响图像级检测和像素级定位之间的权衡[43]-[46]。本文据此考察最大池化与高斯平滑在推理时决策层中的作用。

3. 检测框架

所提方法由三个相互衔接的环节组成: 首先提高异构异常分数的可比性, 其次以适合稀疏缺陷的方式形成图像级分数, 最后简化后处理参数以提供可解释的运行调节。全部改动均位于 MMR 式特征提取器之后, 因此该方法属于推理时决策层, 而非新的异常检测骨干网络。

图 1 给出了完整流程。教师-学生分支与 NFB 分支分别生成蒸馏证据和原型证据; 融合模块先对两类分数进行位置-尺度标准化, 再执行加权组合; 最终通过最大池化获得图像级分数, 并由 PRO-Focus 调节异常图的平滑程度。该流程无需改变骨干网络的训练目标或参数。

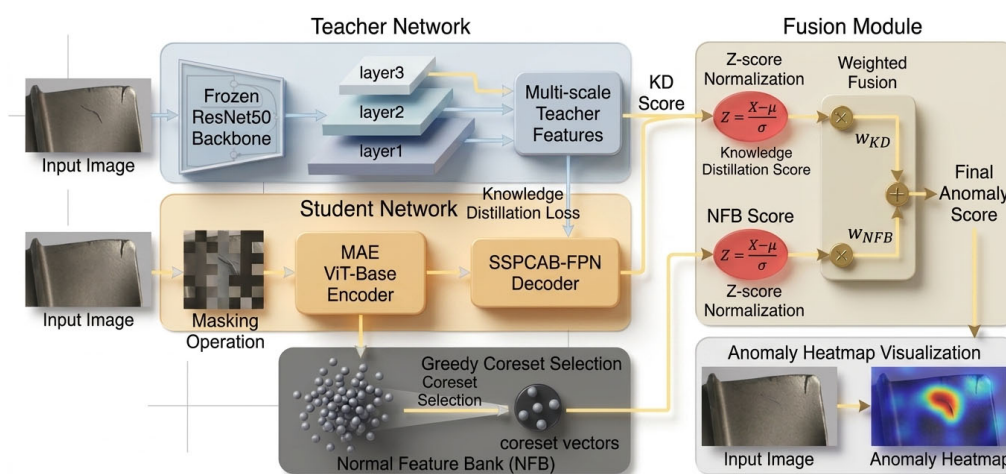


Figure 1. Overview of the DASF pipeline with max-pooling

图 1. 结合最大池化的 DASF 流程概览

3.1. 教师-学生架构

骨干架构沿用 MMR 流程[2]。教师网络为冻结的预训练 ResNet50, 输出三层多尺度特征: layer1 (56×56 , 256 维)、layer2 (28×28 , 512 维)和 layer3 (14×14 , 1024 维)。学生网络由掩码自编码器(Masked Autoencoder, MAE)的 ViT-Base 编码器[47]和采用 SSPCAB 的特征金字塔网络(Feature Pyramid Network, FPN)解码器组成。SSPCAB 即压缩-空间金字塔上下文聚合模块, 用于增强局部结构与边界信息。训练时采用掩码比例 $\eta = 0.4$ 的特征重建, 使学生网络利用上下文恢复正常结构; 计算损失前, 各尺度特征通过最近邻插值对齐至相同空间分辨率。

3.2. 分布对齐分数融合(DASF)

DASF 是附加于固定教师-学生骨干之后的推理时决策层。对于测试图像 x , 它联合两类互补证据: 教师-学生特征差异产生的蒸馏分数 S_{KD} , 以及测试特征相对正常原型库的匹配分数 S_{NFB} 。具体过程如下。

步骤 1: 生成异构分数。蒸馏分数定义为:

$$S_{\text{KD}}(x) = d_{\cos}(f_T(x), f_S(x)), \quad (1)$$

其中, $f_T(\cdot)$ 和 $f_S(\cdot)$ 分别表示教师网络与学生网络的特征表征, $d_{\cos}(\cdot)$ 表示余弦距离。原型匹配分数定义为:

$$S_{\text{NFB}}(x) = \min_{p \in \mathcal{B}_{\text{NFB}}} d(f_T(x), p), \quad (2)$$

其中, $\mathcal{B}_{\text{NFB}}$ 为正常原型库, $d(\cdot)$ 为特征空间中的距离函数。因而, S_{KD} 衡量教师 - 学生表征的不一致程度, S_{NFB} 则衡量测试特征偏离正常特征集合的程度。

步骤 2: 利用正常训练统计量进行位置 - 尺度对齐。由于 S_{KD} 与 S_{NFB} 的数值范围和经验分布不同, 直接加权可能使高幅值分数源占据主导。为此, DASF 在正常训练集 $\mathcal{D}_{\text{train}}^N$ 上分别估计两类分数的均值与标准差:

$$(\mu_{\text{KD}}, \sigma_{\text{KD}}) = \text{MeanStd}\{S_{\text{KD}}(x_i) : x_i \in \mathcal{D}_{\text{train}}^N\}, (\mu_{\text{NFB}}, \sigma_{\text{NFB}}) = \text{MeanStd}\{S_{\text{NFB}}(x_i) : x_i \in \mathcal{D}_{\text{train}}^N\}. \quad (3)$$

推理时, 使用固定的训练集统计量将两类分数映射到标准化尺度:

$$\tilde{S}_{\text{KD}} = \frac{S_{\text{KD}} - \mu_{\text{KD}}}{\sigma_{\text{KD}}}, \tilde{S}_{\text{NFB}} = \frac{S_{\text{NFB}} - \mu_{\text{NFB}}}{\sigma_{\text{NFB}}} \quad (4)$$

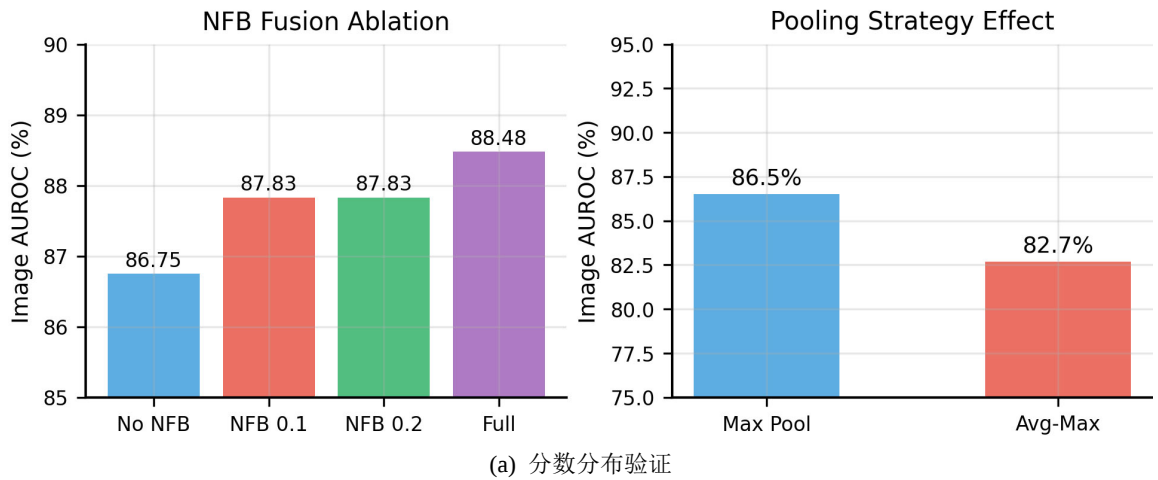
步骤 3: 加权融合。标准化后的分数按下式组合:

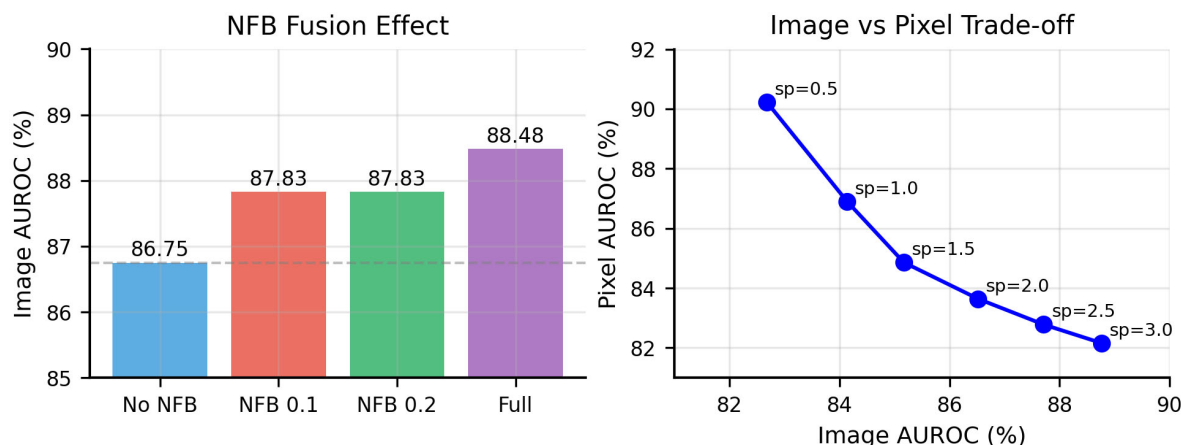
$$S_{\text{final}} = (1 - \alpha) \cdot \tilde{S}_{\text{KD}} + \alpha \cdot \tilde{S}_{\text{NFB}} \quad (5)$$

其中, $\alpha \in [0, 1]$ 为 NFB 分数的融合权重。本文依据消融实验选取 $\alpha = 0.2$, 相关结果见第 8 节。

Z-score 标准化不要求原始分数服从高斯分布, 也不会改变其偏度或尾部形态; 它仅完成位置与尺度变换。该方法计算简单, 但均值和标准差可能受离群值及重尾分布影响。本文采用它的目的, 是以较低推理开销减轻两个分数源之间的尺度支配, 而不是获得概率意义上的校准输出。

图 2 展示了两类训练分数的经验分布。 S_{NFB} 在训练数据上未拒绝正态性假设(显著性水平 0.05, KS 统计量 = 0.0158, $p = 0.69$), 而 S_{KD} 呈现更明显的重尾特征。标准化将两类分数的经验均值和标准差调整至约 0 和 1, 同时保留各自的分布形态。因此, 该图说明的是分数可比性得到改善, 而非两类分数被转换为相同的高斯分布。





(b) DASF 对齐原理

Figure 2. Score-distribution validation and DASF alignment**图 2.** 分数分布验证与 DASF 对齐

表 2 概述了 DASF 的处理过程:

Table 2. Distribution-Aligned Score Fusion (DASF) algorithm**表 2.** 分布对齐分数融合(DASF)算法

算法: 分布对齐分数融合

输入: 训练集特征 $\mathcal{F}_{\text{train}}$, 融合权重 α

输出: 最终异常分数 S_{final}

1. //计算训练集统计量
2. $\mu_{\text{KD}}, \sigma_{\text{KD}} \leftarrow \text{MeanStd}(\{S_{\text{KD}}(x) : x \in \mathcal{F}_{\text{train}}\})$
3. $\mu_{\text{NFB}}, \sigma_{\text{NFB}} \leftarrow \text{MeanStd}(\{S_{\text{NFB}}(x) : x \in \mathcal{F}_{\text{train}}\})$
4. //对测试图像执行推理
5. $S_{\text{KD}} \leftarrow \text{ComputeDistillationScore}(x)$
6. $S_{\text{NFB}} \leftarrow \text{ComputePrototypeScore}(x)$
7. //Z-score 归一化
8. $\tilde{S}_{\text{KD}} \leftarrow (S_{\text{KD}} - \mu_{\text{KD}}) / \sigma_{\text{KD}}$
9. $\tilde{S}_{\text{NFB}} \leftarrow (S_{\text{NFB}} - \mu_{\text{NFB}}) / \sigma_{\text{NFB}}$
10. //加权融合
11. $S_{\text{final}} \leftarrow (1 - \alpha) \cdot \tilde{S}_{\text{KD}} + \alpha \cdot \tilde{S}_{\text{NFB}}$
12. 返回 S_{final}

3.3. 正常特征库(NFB)的构建

正常特征库存储具有代表性的正常特征, 用于原型匹配。其构建采用基于特征空间覆盖的贪心核心集选择思想[48]:

- **特征提取:** 对每幅训练图像提取多尺度教师特征(layer1、layer2、layer3), 随后通过全局平均池化

生成特征向量。

- **贪心核心集选择:** 每个尺度设置 K 个原型的预算(默认 $K = 200$), 以均值特征向量为起点, 根据候选特征与已选原型之间的最大距离迭代选择特征。该策略在控制存储开销的同时, 能够较好地覆盖特征空间。
- **分尺度 NFB:** 为每个特征尺度分别维护原型库, 共得到 $200 \times 3 = 600$ 个向量。

NFB 在训练阶段离线构建, 并在推理期间保持固定; 运行时的额外计算开销仅来自最近原型检索。

该算法使融合步骤明确且可审计: 异构异常证据在组合前先进行标准化, 而不是依据原始分数幅值直接合并。

3.4. 最大池化为何更适合图像级评分

工业缺陷往往只占图像面积的 1%~10%, 由此形成本文关于图像级聚合的基本假设: 平均池化会按照异常区域比例稀释局部响应, 而最大池化能够保留峰值证据, 但也更容易放大正常背景中的极端高分。因此, 最大池化的有效性不仅取决于缺陷面积, 还取决于异常分数与域偏移所引起正常高分的上尾可分性。

设 $S_{\text{pixel}}(i, j)$ 表示位置 (i, j) 的像素级异常分数, 图像含 n 个像素, 异常区域比例为 $\rho \in (0, 1)$ 。为避免与融合权重 α 混淆, 以下以 ρ 表示空间占比。两种聚合方式为:

- **平均池化:** $S_{\text{avg}} = \frac{1}{n} \sum_{i,j} S_{\text{pixel}}(i, j) \approx (1-\rho)\mu_{\text{normal}} + \rho\mu_{\text{anomaly}}$
- **最大池化:** $S_{\text{max}} = \max_{i,j} S_{\text{pixel}}(i, j)$

若正常像素和异常像素的期望分数分别为 μ_N 和 μ_A , 则平均池化在异常图像与正常图像之间的期望差为:

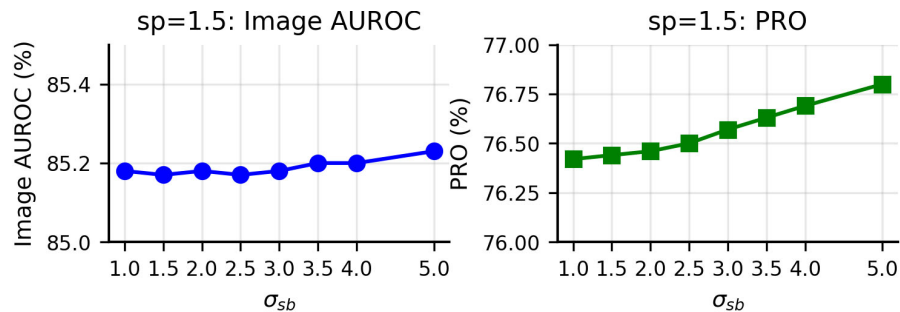
$$\Delta_{\text{avg}} = \rho(\mu_A - \mu_N), \quad (6)$$

当 $\rho \ll 1$ 时, 平均分数的期望差按比例收缩。对于最大池化, 设正常像素与异常像素分数的累积分布函数分别为 $F_N(t)$ 和 $F_A(t)$; 在像素近似独立的简化条件下, 有:

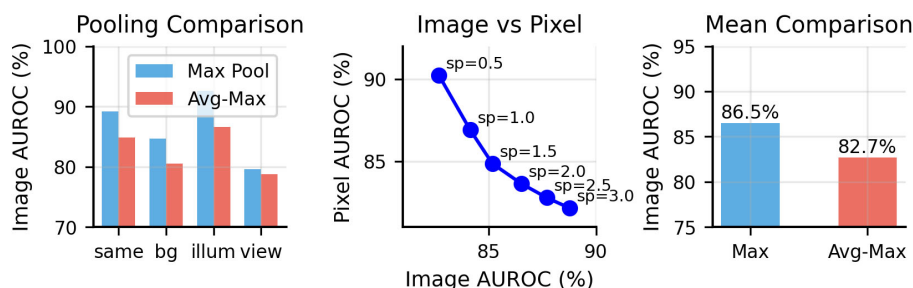
$$\Pr(S_{\text{max}} \leq t | \text{normal}) = F_N(t)^n, \quad \Pr(S_{\text{max}} \leq t | \text{anomaly}) = F_N(t)^{(1-\rho)n} F_A(t)^{\rho n}. \quad (7)$$

当相关阈值处满足 $F_A(t) < F_N(t)$, 即异常分数在上尾具有更高取值概率时, 最大池化更容易保留稀疏异常证据; 若域偏移产生大量正常高分, 则最大池化也可能增加误报警。式(6)因而解释了平均池化的稀释机制, 而最大池化是否适用仍需通过受控实验判断。该分析与极值理论(Extreme Value Theory, EVT)中对尾部事件和样本极值的建模思想一致。

图 3 并列展示参数约化与池化机制: 左侧验证高斯卷积关系, 右侧比较不同聚合策略在同域和域偏移条件下的图像级表现。



(a) 高斯卷积验证



(b) 最大池化机制与消融

Figure 3. Gaussian convolution and max-pooling analysis**图 3.** 高斯卷积与最大池化分析

表 3 给出逐步配置结果。第 1 行至第 2 行同时引入 NFB 并将平均池化替换为最大池化, 因此 10.63 个百分点表示联合配置增益, 不能归因于单一组件。池化策略的直接证据来自图 3 中控制其余设置的比较。第 2 行至第 3 行仅改变 σ , 图像级 AUROC 提高 0.65 个百分点; 完整流程相对“仅蒸馏 + 平均池化”基线的总差值为 11.28 个百分点。

Table 3. Staged comparison of decision-layer configurations on AeBAD-S**表 3.** AeBAD-S 决策层配置的逐步比较

配置	池化方式	σ	图像级 AUROC
仅蒸馏	平均	1.5	77.20%
最大池化 + NFB	最大	1.5	87.83%
最大池化 + NFB	最大	3.0	88.48%
Δ (NFB 与池化联合配置, $\sigma=1.5$)		+10.63%	
Δ ($\sigma=3.0$ 与 $\sigma=1.5$ 比较)		+0.65%	
Δ (完整流程与仅蒸馏比较)		+11.28%	

直接池化比较和逐步配置结果均支持在本研究的域偏移筛查场景中采用最大池化, 但该结论不外推至所有异常检测模型和数据集。

3.5. PRO-Focus 细化算子与参数敏感性

对于取值范围为 $[0,1]$ 、空间尺寸为 $H \times W$ 的原始异常图 A , PRO-Focus 定义为:

$$A' = G_{\sigma_{sb}} \left(G_{\sigma_{sp}} (A) \right) + w \cdot (A - G_{\sigma_{sb}} (A)) \quad (8)$$

$G_{\sigma}(\cdot)$ 表示标准差为 σ 的高斯平滑, σ_{sp} 和 σ_{sb} 分别为预平滑与边界平滑参数, w 为边界增强权重。

三个耦合参数增加了部署调节的复杂度。连续高斯卷积可等价一次具有合成标准差的高斯卷积, 这一性质可用于约化两个平滑参数。

定理 1 (高斯卷积可加性)。 对于方差为 σ^2 的高斯核 G_{σ} , 有:

$$G_{\sigma_1} * G_{\sigma_2} = G_{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (9)$$

证明。 利用傅里叶变换性质 $\mathcal{F}\{G_{\sigma}\}(\omega) = \exp(-\sigma^2 \omega^2 / 2)$, 可得:

$$\begin{aligned}
\mathcal{F}\{G_{\sigma_1} * G_{\sigma_2}\}(\omega) &= \mathcal{F}\{G_{\sigma_1}\}(\omega) \cdot \mathcal{F}\{G_{\sigma_2}\}(\omega) \\
&= \exp(-\sigma_1^2 \omega^2 / 2) \cdot \exp(-\sigma_2^2 \omega^2 / 2) \\
&= \exp(-(\sigma_1^2 + \sigma_2^2) \omega^2 / 2) \\
&= \mathcal{F}\{G_{\sqrt{\sigma_1^2 + \sigma_2^2}}\}(\omega)
\end{aligned} \tag{10}$$

进行傅里叶逆变换即可得到式(9)。

令 $\sigma_{sp} = \sigma_{sb} = \sigma$ ，式(8)可写为：

$$A' = G_{\sigma\sqrt{2}}(A) + w \cdot (A - G_{\sigma}(A)) \tag{11}$$

上述恒等式从数学上合并了两个高斯参数；对 w 的处理则来自经验结果。将 w 从 0.25 调整为 0.0 时，各项性能差值不足 0.15 个百分点，因此后续实验固定 w ，仅以 σ 作为主要运行参数。换言之，单参数化同时依赖高斯卷积关系和对 w 的经验敏感性分析。

在 $\sigma_{sp} = \sigma_{sb}$ 的设定下，图像级检测与像素级定位之间的权衡可由单一参数 σ 描述：

- $\sigma = 1.5$ ：平滑程度最低，保留边界细节→图像级 87.43%/像素级 83.90%/PRO76.03%。
- $\sigma = 2.0$ ：取得均衡权衡→图像级 87.93%/像素级 83.12%/PRO73.22%。
- $\sigma = 2.5$ ：采用更强平滑→图像级 87.72%/像素级 82.78%/PRO69.92%。
- $\sigma = 3.0$ ：平滑程度最高→图像级 88.48%/像素级 82.14%/PRO73.95%。

较小的 σ 保留更多局部与边界信息，较大的 σ 更强调平滑后的图像级响应。部署人员可根据筛查或定位优先级选择运行点，无需重新训练骨干网络。

图 4 进一步汇总单参数平滑的运行谱及各测试条件下的性能差异。

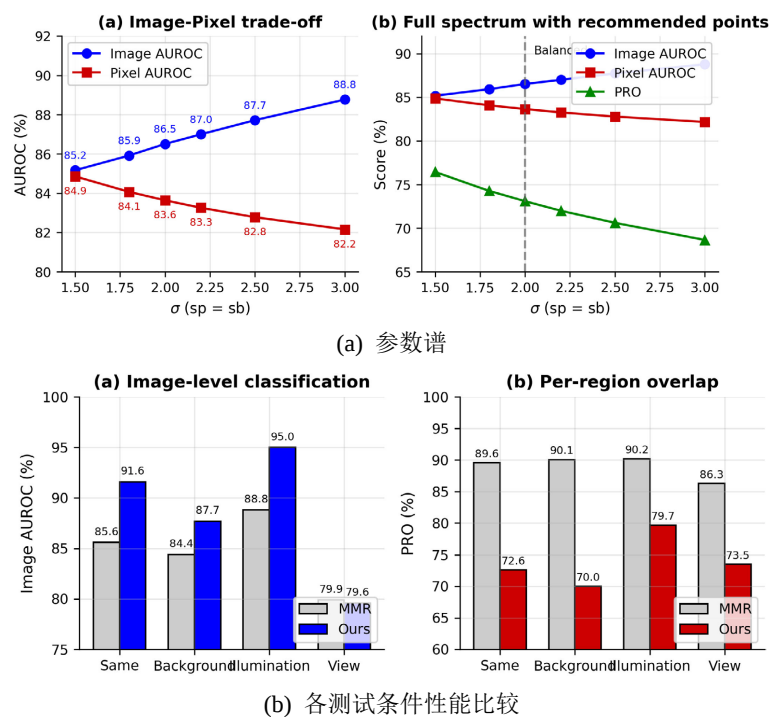


Figure 4. Parameter sensitivity and per-domain performance

图 4. 参数敏感性与分条件性能

图 4(a)显示, 随着 σ 增大, 图像级 AUROC 总体上升, 而像素级 AUROC 与 PRO 总体下降; 图 4(b)显示, 相对 MMR 的差值随测试条件而变化, 本文方法并非在所有条件下均占优。这两点分别构成 RQ3 参数敏感性分析和分条件报告的依据。

4. 实验设置

4.1. 实验设计与研究问题

实验围绕三个研究问题组织, 每个问题对应决策层中的一个组件, 并通过 AeBAD-S 上的对照实验进行评估。

- **RQ1: 经位置 - 尺度标准化的 NFB 证据能否改善融合结果?** 在骨干网络、池化和平滑设置不变时, 以 $\alpha = 0$ 作为不使用 NFB 的对照, 并改变 NFB 融合权重。该实验评价 DASF 融合分支的增量价值, 不构成 Z-score 与其他标准化方法的比较。
- **RQ2: 最大池化能否改善稀疏缺陷的图像级聚合?** 通过控制其余设置比较最大池化与 avg_max, 并以逐步配置结果补充端到端表现。池化的直接差值与完整流程的联合增益分别报告, 以避免混合归因。
- **RQ3: 单参数平滑能否控制图像级检测与像素级定位之间的权衡?** 令 $\sigma_{sp} = \sigma_{sb} = \sigma$, 改变 σ 并同时报告图像级 AUROC、像素级 AUROC 和 PRO, 以检验其作为部署运行参数的有效性。

上述设计遵循“研究问题 - 对照条件 - 评价指标”的对应关系: RQ1 评价融合证据, RQ2 评价聚合方式, RQ3 评价运行参数。除逐步配置实验外, 单组件结论均来自其余设置不变的对照。由于未报告重复运行和置信区间, 本文仅描述观测差值, 不作统计显著性判断。

4.2. 数据集

主要评估基准为 AeBAD-S (航空发动机叶片异常检测 - 单叶片) [2]。正常训练集包含 521 幅图像; 评估部分包含 1,639 幅图像, 覆盖同域条件及三类领域偏移:

- **同域(Same):** 测试集正常样本与训练分布一致(689 幅: 230 幅正常图像 + 459 幅异常图像)。
- **背景(Background):** 训练集与测试集的背景纹理不同(305 幅: 93 幅正常图像 + 212 幅异常图像)。
- **光照(Illumination):** 光照条件发生变化(273 幅: 75 幅正常图像 + 198 幅异常图像)。
- **视角(View):** 拍摄位置或视角相对训练条件发生变化(372 幅: 92 幅正常图像 + 280 幅异常图像)。

所有异常图像均提供像素级标注, 缺陷类型包括破损、烧蚀、断裂和沟槽。

本文另在 MVTec AD [3] [5]上开展辅助评估。该数据集包含 15 个工业物体与纹理类别, 共 5,354 幅图像, 缺陷类型包括划痕、凹痕、污染和结构损伤。由于 MVTec AD 实验与 AeBAD-S 使用不同骨干网络和聚合配置, 其结果用于考察跨数据集适用性, 而非严格受控的直接验证。

4.3. 评估指标

遵循工业异常检测的常用评估协议[9], 本文报告三项指标。图像级 AUROC 衡量正常/异常图像排序能力, 图像分数由平均池化、最大池化或 avg_max 获得; 像素级 AUROC 衡量像素层面的异常定位能力; PRO (Per-Region-Overlap)计算预测区域与真实缺陷区域的重叠程度, 并对各连通缺陷区域等权统计。三项指标分别对应产品级判定、像素级排序和区域覆盖质量。

5. 融合与决策聚合评估

5.1. AeBAD-S 上的总体性能

表 4 汇总了 AeBAD-S 上的总体结果。当 $\sigma = 3.0$ 时, 本文方法的图像级 AUROC 为 88.48%, 较 MMR

的 84.7%高 3.78 个百分点; 当 $\sigma=1.5$ 时, 该指标为 87.43%, 差值为 2.73 个百分点。相应地, $\sigma=1.5$ 时的像素级 AUROC 为 83.90%, 较 MMR 低 1.40 个百分点; $\sigma=1.5$ 和 $\sigma=3.0$ 时的 PRO 分别为 76.03%和 73.95%, 均低于 MMR 的 89.1%。这些结果显示图像级排序能力提高, 同时像素级定位尤其是区域覆盖能力下降。

Table 4. Overall results on AeBAD-S

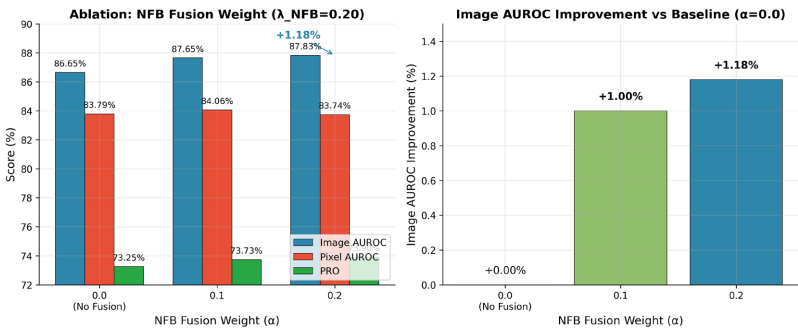
表 4. AeBAD-S 总体结果

配置	图像级 AUROC	像素级 AUROC	PRO	说明
MMR [2]	84.7%	85.3%	89.1%	各测试条件平均
本文方法($\sigma=1.5$)	87.43%	83.90%	76.03%	像素级导向
本文方法($\sigma=3.0$)	88.48%	82.14%	73.95%	图像级导向

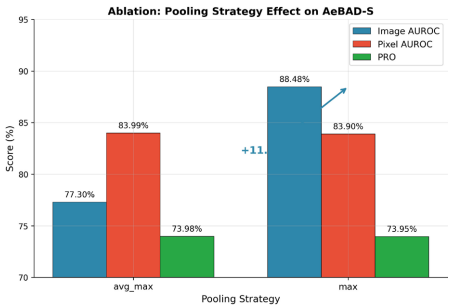
总体结果支持将该流程用于以产品级筛查为首要目标的场景; 若任务更关注缺陷区域的完整覆盖, 则较低的 PRO 构成明确限制。

5.2. 消融实验: NFB 融合权重

如表 5 所示, 将 NFB 融合权重由 0.0 增至 0.2 后, 图像级 AUROC 从 86.65%提高到 87.83%, 像素级 AUROC 则从 83.79%变为 83.74%。在当前配置下, 经标准化的原型证据主要改善图像级排序, 对像素级指标的影响较小。由于未报告重复运行方差或置信区间, 1.18 个百分点仅表示观测差值。图 5 给出了对应可视化结果。



(a) NFB 融合权重



(b) 池化策略

Figure 5. Ablation studies on AeBAD-S

图 5. AeBAD-S 消融实验

Table 5. Ablation of the NFB fusion weight on AeBAD-S
表 5. AeBAD-S 上 NFB 融合权重消融

融合权重(α)	图像级 AUROC	像素级 AUROC	PRO
0.0	86.65%	83.79%	73.25%
0.1	87.65%	84.06%	73.73%
0.2	87.83%	83.74%	73.90%
Δ (0.2 vs 0.0)	+1.18%	-0.05%	+0.65%

NFB 消融结果表明，标准化后的原型证据可为图像级决策提供补充信息，而像素级 AUROC 在当前权重范围内变化较小。

5.3. 与先进方法的比较

表 6 显示，在 $\sigma=3.0$ 时，本文方法的图像级 AUROC 在同域、背景和光照条件下高于 MMR，在视角偏移下为 79.60%，略低于 MMR 的 79.9%。四种测试条件的平均值为 88.48%，较 MMR 高 3.78 个百分点；其中光照偏移下的差值最大(95.00%对 88.8%)。

像素级指标呈现相反趋势：当 $\sigma=1.5$ 时，像素级 AUROC 为 83.90%，较 MMR 的 85.3%低 1.40 个百分点。该结果说明位置 - 尺度融合和面向图像级判定的后处理未同步改善像素排序性能；在缺少统计检验的情况下，本文仅报告差值大小。

PRO 的下降更为明显(表 7)。当 $\sigma=1.5$ 时，本文方法的平均 PRO 为 76.03%，低于 MMR 的 89.1%；在同域条件下(表 8)，两者分别为 76.44%和 89.6%。PRO 关注预测区域与完整缺陷区域的重叠，而本方法的最大池化强调局部峰值能否触发产品级判定，二者优化目标并不一致。因此，本文方法适用于筛查优先的检测流程，而不适合作为追求精确区域覆盖的分割方案。第 9 节进一步讨论这一适用边界。

Table 6. Image AUROC and Pixel AUROC on AeBAD-S
表 6. AeBAD-S 上的图像级 AUROC 与像素级 AUROC

方法	图像级 AUROC (%)					像素级 AUROC (%)				
	同域	背景	光照	视角	平均	同域	背景	光照	视角	平均
PatchCore [17]	75.2	74.1	74.6	60.1	71.0	86.3	83.7	80.9	76.2	81.8
Reverse Distillation [10]	82.4	84.3	85.5	71.9	81.0	87.1	85.6	83.8	79.4	84.0
DRAEM [49]	64.0	62.1	61.6	62.3	62.5	71.2	69.4	65.3	68.1	68.5
NSA [50]	66.5	48.8	55.5	55.9	56.7	72.8	58.6	62.1	60.5	63.5
MMR [2]	85.6	84.4	88.8	79.9	84.7	87.3	85.9	86.1	82.0	85.3
本文方法($\sigma=1.5$)	90.20	86.50	94.20	78.80	87.43	84.86	83.24	85.41	82.09	83.90
本文方法($\sigma=2.0$)	90.80	87.10	94.60	79.20	87.93	83.98	82.31	84.52	81.67	83.12
本文方法($\sigma=3.0$)	91.60	87.70	95.00	79.60	88.48	83.11	81.20	83.67	80.56	82.14

分条件结果表明，本文方法的图像级收益主要来自同域、背景和光照条件，视角偏移下未超过 MMR；像素级 AUROC 则总体低于 MMR。因此，改进应解释为产品级排序能力的提升，而非检测与定位性能的同步提升。

Table 7. PRO on AeBAD-S
表 7. AeBAD-S 上的 PRO

方法	PRO (%)				
	同域	背景	光照	视角	平均
PatchCore [17]	89.5	89.4	88.2	84.0	87.8
ReverseDistillation [10]	86.4	86.4	86.7	82.9	85.6
DRAEM [49]	71.4	44.3	67.6	71.1	63.6
NSA [50]	43.0	29.7	59.9	51.1	45.9
MMR [2]	<u>89.6</u>	<u>90.1</u>	<u>90.2</u>	<u>86.3</u>	<u>89.1</u>
本文方法($\sigma = 1.5$)	76.44	75.23	77.65	74.81	76.03
本文方法($\sigma = 2.0$)	74.19	72.46	74.38	71.83	73.22
本文方法($\sigma = 3.0$)	72.60	70.00	79.70	73.50	73.95

PRO 结果进一步限定了方法用途：它可作为图像级筛查的决策层，但不具备相对 MMR 的区域覆盖优势。

Table 8. Same-domain results on AeBAD-S
表 8. AeBAD-S 同域结果

方法	图像级 AUROC (%)			像素级 AUROC (%)			PRO (%)		
	$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 3.0$	$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 3.0$	$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 3.0$
MMR [2]		85.6			87.3			89.6	
本文方法	90.20	90.80	91.60	84.86	83.98	83.11	76.44	74.19	72.60
Δ		+6.00			-4.19			-16.96	

同域比较表明，图像级 AUROC 与 PRO 之间的权衡在不存在领域偏移时仍然可见，因此该现象不能仅归因于领域偏移；聚合方式与平滑设置也是可能原因。

5.4. MVTec AD 上的辅助结果

为辅助考察跨数据集适用性，本文采用 ab_final_opt 配置在 MVTec AD 上进行逐类别训练：教师网络为 WideResNet50，启用边界损失，focal gamma = 2.0，采用 avg_max 池化，并设置 $\sigma_{sp} = 1.0$ 、 $\sigma_{sb} = 3.5$ 、 $w = 0.5$ 。该配置用于在同域数据上平衡图像级检测与像素级定位。表 9 报告各类别结果。

Table 9. Results on MVTec AD
表 9. MVTec AD 结果

Category	图像级 AUROC	像素级 AUROC	PRO
Bottle	99.9%	96.40%	87.6%
Cable	95.2%	93.16%	74.9%
Capsule	97.6%	97.00%	90.3%

续表

Carpet	100.0%	97.31%	90.7%
Grid	99.9%	98.43%	93.7%
Hazelnut	100.0%	97.95%	89.6%
Leather	99.8%	98.62%	96.8%
Metal Nut	99.7%	95.63%	82.0%
Pill	94.3%	98.27%	88.3%
Screw	90.3%	98.86%	94.1%
Tile	99.0%	91.14%	75.4%
Toothbrush	99.4%	97.46%	89.0%
Transistor	99.5%	93.46%	77.8%
Wood	98.2%	92.53%	81.1%
Zipper	97.3%	95.40%	84.8%
平均	98.01%	96.11%	86.41%

在该配置下，MVTecAD 的平均图像级 AUROC、像素级 AUROC 和 PRO 分别为 98.01%、96.11%和 86.41%，说明 avg_max 在该同域数据集上取得了较均衡的三项指标。

将 MVTecAD 上的聚合方式改为最大池化后，平均图像级 AUROC 为 98.31%，较 avg_max 高 0.30 个百分点，同时像素级 AUROC 和 PRO 略有下降。该结果与 AeBAD-S 上完整流程相对基线的 11.28 个百分点差值不可直接等同，因为后者还包含 NFB 与平滑设置的变化。两组结果共同说明，聚合策略的收益依赖数据条件和评价目标。

需要强调的是，MVTec AD 与 AeBAD-S 实验分别使用 WideResNet50 和 ResNet50，默认聚合方式也不同。因此，MVTec AD 结果只能提供辅助的外部有效性证据，不能视为完全相同配置下的跨数据集受控验证。

6. 实践讨论

6.1. 实践价值

所提决策层位于特征提取之后，不改变教师 - 学生网络的训练目标，因而可作为现有检测流程的增量升级。对于重新训练需要重新采集正常样本、验证生产约束或中断既有流程的场景，这种部署方式能够降低改造成本。NFB 权重消融中的 1.18 个百分点图像级差值虽小于完整流程增益，但说明原型证据可以在保持骨干网络不变的情况下提供补充信息。

6.2. 合格/不合格筛查的工程权衡

最大池化将局部强响应视为产品级异常证据，这与“一个关键缺陷即可判定零件不合格”的质量控制规则一致。实验中图像级 AUROC 提高而 PRO 下降，说明峰值驱动的筛查与完整区域覆盖之间存在明确权衡。对于先自动标记可疑零件、再由人工核验缺陷位置的流程，该运行点具有实际意义；对于需要自动测量缺陷面积、形状或边界的系统，则应优先采用较小的 σ 、更均衡的聚合方式，或增加专门的边界细化模块。

6.3. 单参数平滑带来的运行简洁性

将两个高斯参数设为统一的 σ ，并根据敏感性实验固定 w ，可将部署调节从三维参数组合简化为一维运行选择。若优先考虑图像级筛查，可采用 $\sigma=3.0$ ；若需要保留更多局部与边界信息，可采用 $\sigma=1.5$ 。由于这一调整仅发生在推理阶段，同一骨干模型可在不同生产要求下切换运行点。不过，当前推荐值来自有限数据集，跨生产线使用前仍需在本地验证集上复核。

7. 局限性

7.1. NFB 的计算开销

NFB 为三个尺度各保留 $K=200$ 个原型，共存储 600 个向量；推理时需在各尺度执行最近原型检索。在 NVIDIA RTX 3090 上，该步骤使每幅图像的延迟增加约 15~20 ms。对于延迟敏感型应用，可减小 K ，或采用近似最近邻检索与向量索引。

7.2. 均值与标准差估计的敏感性

Z-score 标准化不要求分数服从高斯分布，但均值与标准差对重尾样本和离群值较敏感。当部署域中的正常样本产生异常高分时，训练统计量可能无法提供稳定尺度。后续可比较中位数 - 绝对中位差、 S_n/Q_n 等稳健尺度估计、截尾统计量和百分位排序，以评估更稳健的对齐方式[51]。

7.3. 对严重领域偏移的泛化能力

AeBAD-S 实验覆盖同域、背景、光照和视角四种测试条件，但尚未覆盖传感器模态变化、全新纹理或更严重的复合偏移。最优融合权重 α 可能随偏移类型和强度变化，固定权重在极端部署条件下未必适用。未来可根据正常分数分布的变化幅度自适应选择 α 。

7.4. 统计不确定性与受控比较

本文报告的是单次实验结果，尚未提供多随机种子重复、置信区间或显著性检验，因此较小的指标差值只能解释为观测结果。与此同时，MVTec AD 与 AeBAD-S 使用不同骨干网络和默认聚合配置，跨数据集结果不构成严格受控比较。后续研究应在统一配置下进行重复实验，并报告不确定性区间与计算成本。

8. 结论

本文针对领域偏移下的产品级判定问题，提出了附加于固定教师 - 学生骨干之后的 DASF 后处理领域自适应模块。这里的“分布对齐”具体指利用正常训练集的一阶矩(均值)和二阶中心矩(方差/标准差)，对蒸馏分数与原型分数实施位置 - 尺度标准化；该操作用于提高异构分数的可比性，不涉及高阶矩匹配或完整概率分布对齐。模块进一步以最大池化保留稀疏异常响应，并将后处理简化为以 σ 为主的运行调节。

AeBAD-S 结果表明，完整配置的平均图像级 AUROC 为 88.48%，较 MMR 高 3.78 个百分点；相应的像素级 AUROC 与 PRO 均有所下降。这一结果将方法的适用范围限定为产品级筛查优先、允许后续人工核验的场景，而非精确区域分割。MVTecAD 的辅助结果进一步说明，最大池化的收益依赖数据条件和评价目标。

总体而言，实验支持将 DASF 视为一种无需目标域标签、无需修改或重新训练源模型的轻量化后处理适应方案：它通过一阶与二阶统计量标准化、原型证据融合及峰值聚合，改善检测器在领域偏移条件

下的产品级排序表现。该结论仅表明后处理设计能够增强特定教师-学生流程的部署泛化性, 并不意味着模型已经学习到目标域分布, 也不代表像素级定位能力同步提升。由于当前证据仍受数据集范围、配置差异和统计不确定性限制, 后续需在统一配置、更多偏移类型及不同传感模态下进行验证。

参考文献

- [1] Cao, T., Zhu, J. and Pang, G. (2023) Anomaly Detection under Distribution Shift. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 6488-6500. <https://doi.org/10.1109/iccv51070.2023.00599>
- [2] Zhang, Z., Zhao, Z., Zhang, X., Sun, C. and Chen, X. (2023) Industrial Anomaly Detection with Domain Shift: A Real-World Dataset and Masked Multi-Scale Reconstruction. *Computers in Industry*, **151**, Article 103990. <https://doi.org/10.1016/j.compind.2023.103990>
- [3] Bergmann, P., Fauser, M., Sattlegger, D. and Steger, C. (2019) MVTec AD—A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 9584-9592. <https://doi.org/10.1109/cvpr.2019.00982>
- [4] Bergmann, P., Jin, X., Sattlegger, D. and Steger, C. (2022) The MVTec 3D-AD Dataset for Unsupervised 3D Anomaly Detection and Localization. *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, **5**, 202-213. <https://doi.org/10.5220/0010865000003124>
- [5] Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D. and Steger, C. (2022) Beyond Dents and Scratches: Logical Constraints in Unsupervised Anomaly Detection and Localization. *International Journal of Computer Vision*, **130**, 947-969. <https://doi.org/10.1007/s11263-022-01578-9>
- [6] Ben Ammar, M., Mendoza, A., Belkhir, N., Manzanera, A. and Franchi, G. (2025) Foundation Models and Transformers for Anomaly Detection: A Survey. *Information Fusion*, **126**, Article 103517.
- [7] Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., et al. (2024) Deep Industrial Image Anomaly Detection: A Survey. *Machine Intelligence Research*, **21**, 104-135. <https://doi.org/10.1007/s11633-023-1459-z>
- [8] Batzner, K., Heckler, L. and König, R. (2024) EfficientAD: Accurate Visual Anomaly Detection at Millisecond-Level Latencies. 2024 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 3-8 January 2024, 127-137. <https://doi.org/10.1109/wacv57701.2024.00020>
- [9] Bergmann, P., Fauser, M., Sattlegger, D. and Steger, C. (2020) Uninformed Students: Student-Teacher Anomaly Detection with Discriminative Latent Embeddings. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 4182-4191. <https://doi.org/10.1109/cvpr42600.2020.00424>
- [10] Deng, H. and Li, X. (2022) Anomaly Detection via Reverse Distillation from One-Class Embedding. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 9727-9736. <https://doi.org/10.1109/cvpr52688.2022.00951>
- [11] Tang, C., Zhou, S., Li, Y., Dong, Y. and Wang, L. (2026) Advancing Pre-Trained Teacher: Towards Robust Feature Discrepancy for Anomaly Detection. *IEEE Transactions on Image Processing*, **35**, 4509-4522. <https://doi.org/10.1109/TIP.2026.3683277>
- [12] Salehi, M., Sadjadi, N., Baselizadeh, S., Rohban, M.H. and Rabiee, H.R. (2021) Notice of Removal: Multiresolution Knowledge Distillation for Anomaly Detection. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 14897-14907. <https://doi.org/10.1109/cvpr46437.2021.01466>
- [13] Wang, G., Han, S., Ding, E. and Huang, D. (2021) Student-Teacher Feature Pyramid Matching for Anomaly Detection. *Proceedings of the British Machine Vision Conference 2021*, Online, 22-25 November 2021, 1-14. <https://doi.org/10.5244/c.35.349>
- [14] Xu, J. and Jiang, S. (2025) Hierarchical Knowledge Transfer: Cross-Layer Distillation for Industrial Anomaly Detection. *Journal of Imaging*, **11**, Article 102. <https://doi.org/10.3390/jimaging11040102>
- [15] Cohen, N. and Hoshen, Y. (2020) Sub-Image Anomaly Detection with Deep Pyramid Correspondences. <https://arxiv.org/pdf/2005.02357>
- [16] Defard, T., Setkov, A., Loesch, A. and Audigier, R. (2021) PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization. In: Del Bimbo, A., et al., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 475-489. https://doi.org/10.1007/978-3-030-68799-1_35
- [17] Roth, K., Pemula, L., Zepeda, J., Scholkopf, B., Brox, T. and Gehler, P. (2022) Towards Total Recall in Industrial Anomaly Detection. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 14298-14308. <https://doi.org/10.1109/cvpr52688.2022.01392>
- [18] Fang, Z., Wang, X., Li, H., Liu, J., Hu, Q. and Xiao, J. (2023) FastRecon: Few-Shot Industrial Anomaly Detection via

- Fast Feature Reconstruction. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 17435-17444. <https://doi.org/10.1109/iccv51070.2023.01603>
- [19] Hu, J., Chen, X., Gan, Z., *et al.* (2024) DMAD: Dual Memory Bank for Real-World Anomaly Detection. arXiv:2403.12362.
- [20] Lee, S., Lee, S. and Song, B.C. (2022) CFA: Coupled-Hypersphere-Based Feature Adaptation for Target-Oriented Anomaly Localization. *IEEE Access*, **10**, 78446-78454. <https://doi.org/10.1109/access.2022.3193699>
- [21] Liang, Y., Zhang, J., Zhao, S., Wu, R., Liu, Y. and Pan, S. (2023) Omni-Frequency Channel-Selection Representations for Unsupervised Anomaly Detection. *IEEE Transactions on Image Processing*, **32**, 4327-4340. <https://doi.org/10.1109/tip.2023.3293772>
- [22] Dai, S., Wu, Y., Li, X. and Xue, X. (2024) Generating and Reweighting Dense Contrastive Patterns for Unsupervised Anomaly Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**, 1454-1462. <https://doi.org/10.1609/aaai.v38i2.27910>
- [23] You, Z., Cui, L., Shen, Y., Yang, K., Lu, X., Zheng, Y., *et al.* (2022) A Unified Model for Multi-Class Anomaly Detection. *Advances in Neural Information Processing Systems* 35, New Orleans, 28 November-9 December 2022, 4571-4584. <https://doi.org/10.52202/068431-0330>
- [24] Jeong, J., Zou, Y., Kim, T., Zhang, D., Ravichandran, A. and Dabeer, O. (2023) WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 19606-19616. <https://doi.org/10.1109/cvpr52729.2023.01878>
- [25] Zhou, Q., Pang, G., Tian, Y., He, S. and Chen, J. (2024) AnomalyCLIP: Object-Agnostic Prompt Learning for Zero-Shot Anomaly Detection. *Proceedings of the Twelfth International Conference on Learning Representations*, Vienna, 7-11 May 2024, 1-33.
- [26] Li, X., Zhang, Z., Tan, X., Chen, C., Qu, Y., Xie, Y., *et al.* (2024) PromptAD: Learning Prompts with Only Normal Samples for Few-Shot Anomaly Detection. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 16848-16858. <https://doi.org/10.1109/cvpr52733.2024.01594>
- [27] Lei, T., Wang, B., Chen, S., Cao, S. and Zou, N. (2024) Texture-AD: An Anomaly Detection Dataset and Benchmark for Real Algorithm Development. arXiv:2409.06367.
- [28] Wang, C., Zhu, W., Gao, B., Gan, Z., Zhang, J., Gu, Z., *et al.* (2024) Real-IAD: A Real-World Multi-View Dataset for Benchmarking Versatile Industrial Anomaly Detection. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 22883-22892. <https://doi.org/10.1109/cvpr52733.2024.02159>
- [29] Xie, G., Wang, J., Liu, J., Lyu, J., Liu, Y., Wang, C., *et al.* (2024) IM-IAD: Industrial Image Anomaly Detection Benchmark in Manufacturing. *IEEE Transactions on Cybernetics*, **54**, 2720-2733. <https://doi.org/10.1109/tcyb.2024.3357213>
- [30] Zou, Y., Jeong, J., Pemula, L., Zhang, D. and Dabeer, O. (2022) Spot-the-Difference Self-Supervised Pre-Training for Anomaly Detection and Segmentation. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M. and Hassner, T., Eds., *Lecture Notes in Computer Science*, Springer, 392-408. https://doi.org/10.1007/978-3-031-20056-4_23
- [31] Kull, M., Silva Filho, T.M. and Flach, P. (2017) Beyond Sigmoids: How to Obtain Well-Calibrated Probabilities from Binary Classifiers with Beta Calibration. *Electronic Journal of Statistics*, **11**, 5052-5080. <https://doi.org/10.1214/17-ejs1338si>
- [32] Platt, J.C. (1999) Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. MIT Press.
- [33] Zadrozny, B. and Elkan, C. (2002) Transforming Classifier Scores into Accurate Multiclass Probability Estimates. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton Alberta, 23-26 July 2002, 694-699. <https://doi.org/10.1145/775047.775151>
- [34] Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning*. 2nd Edition, Springer.
- [35] Kriegel, H., Kroger, P., Schubert, E. and Zimek, A. (2011) Interpreting and Unifying Outlier Scores. *Proceedings of the 2011 SIAM International Conference on Data Mining*, Mesa, 28-30 April 2011, 13-24. <https://doi.org/10.1137/1.9781611972818.2>
- [36] Castanedo, F. (2013) A Review of Data Fusion Techniques. *The Scientific World Journal*, **2013**, Article ID: 704504. <https://doi.org/10.1155/2013/704504>
- [37] Hall, D.L. and Llinas, J. (1997) An Introduction to Multisensor Data Fusion. *Proceedings of the IEEE*, **85**, 6-23. <https://doi.org/10.1109/5.554205>
- [38] Kittler, J., Hatef, M., Duin, R.P.W. and Matas, J. (1998) On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 226-239. <https://doi.org/10.1109/34.667881>
- [39] Costanzino, A., Ramirez, P.Z., Lisanti, G. and Di Stefano, L. (2024) Multimodal Industrial Anomaly Detection by Cross-modal Feature Mapping. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-

- 22 June 2024, 17234-17243. <https://doi.org/10.1109/cvpr52733.2024.01631>
- [40] Tao, C., Cao, X. and Du, J. (2025) G²SF: Geometry-Guided Score Fusion for Multimodal Industrial Anomaly Detection. 2025 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Honolulu, 19-25 October 2025, 20551-20560. <https://doi.org/10.1109/iccv51701.2025.01911>
 - [41] Chae, J., Luo, L., Liu, Y., Wang, R., Yu, D., Liang, Z., *et al.* (2026) StructCore: Structure-Aware Image-Level Scoring for Training-Free Unsupervised Anomaly Detection. arXiv:2602.17048v1.
 - [42] de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory: An Introduction*. Springer.
 - [43] Ardelean, A. and Weyrich, T. (2024) High-Fidelity Zero-Shot Texture Anomaly Localization Using Feature Correspondence Analysis. 2024 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 3-8 January 2024, 1123-1133. <https://doi.org/10.1109/wacv57701.2024.00117>
 - [44] Huang, D., Xue, F., Li, X. and Zhou, Y. (2025) AnoRefiner: Anomaly-Aware Group-Wise Refinement for Zero-Shot Industrial Anomaly Detection. arXiv:2511.22595.
 - [45] Soni, J., Prabakar, N. and Upadhyay, H. (2022) Ada-Thres: An Adaptive Thresholding Method to Mitigate the False Alarms. 2022 *International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, 14-16 December 2022, 916-921. <https://doi.org/10.1109/csci58124.2022.00164>
 - [46] Wang, M., Ding, H., Liew, J.H., Liu, J., Zhao, Y. and Wei, Y. (2023) SegRefiner: Towards Model-Agnostic Segmentation Refinement with Discrete Diffusion Process. *Advances in Neural Information Processing Systems* 36, New Orleans, 10-16 December 2023, 79761-79780. <https://doi.org/10.52202/075280-3492>
 - [47] He, K., Chen, X., Xie, S., Li, Y., Dollar, P. and Girshick, R. (2022) Masked Autoencoders Are Scalable Vision Learners. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 15979-15988. <https://doi.org/10.1109/cvpr52688.2022.01553>
 - [48] Sener, O. and Savarese, S. (2018) Active Learning for Convolutional Neural Networks: A Core-Set Approach. <https://arxiv.org/abs/1708.00489>
 - [49] Zavrtanik, V., Kristan, M. and Skocaj, D. (2021) DRÆM—A Discriminatively Trained Reconstruction Embedding for Surface Anomaly Detection. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 8310-8319. <https://doi.org/10.1109/iccv48922.2021.00822>
 - [50] Schlüter, H.M., Tan, J., Hou, B. and Kainz, B. (2022) Natural Synthetic Anomalies for Self-Supervised Anomaly Detection and Localization. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M. and Hassner, T., Eds., *Lecture Notes in Computer Science*, 474-489. https://doi.org/10.1007/978-3-031-19821-2_27
 - [51] Rousseeuw, P.J. and Croux, C. (1993) Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, **88**, 1273-1283. <https://doi.org/10.1080/01621459.1993.10476408>