

# CFA: 基于耦合超球面的特征自适应面向目标异常定位

李成旭 仁荷大学  
韩国仁川

lsw2646@gmail.com

李承贤 仁荷大学  
韩国仁川

lsh910703@gmail.com

Byung Cheol Song  
仁荷大学 韩国仁川

bcsong@inha.ac.kr

## 摘要

*For a long time, anomaly localization has been widely used in industries. Previous studies focused on approximating the distribution of normal features without adaptation to a target dataset. However, since anomaly localization should precisely discriminate normal and abnormal features, the absence of adaptation may make the normality of abnormal features overestimated. Thus, we propose Coupled-hypersphere-based Feature Adaptation (CFA) which accomplishes sophisticated anomaly localization using features adapted to the target dataset. CFA consists of (1) a learnable patch descriptor that learns and embeds target-oriented features and (2) scalable memory bank independent of the size of the target dataset. And, CFA adopts transfer learning to increase the normal feature density so that abnormal features can be clearly distinguished by applying patch descriptor and memory bank to a pre-trained CNN. The proposed method outperforms the previous methods quantitatively and qualitatively. For example, it provides an AUROC score of 99.5% in anomaly detection and 98.5% in anomaly localization of MVTec AD benchmark. In addition, this paper points out the negative effects of biased features of pre-trained CNNs and emphasizes the importance of the adaptation to the target dataset. The code is publicly available at [https://github.com/sungwool/CFA\\_for\\_anomaly\\_localization](https://github.com/sungwool/CFA_for_anomaly_localization)*

## 1. 引言

异常检测是一项著名的计算机视觉任务，旨在检测给定图像中的异常特征。人类视觉系统（HVS）能够轻松识别图像中的意外模式（即异常），无论特征复杂度如何。随着卷积神经网络（CNN）的快速发展，机器视觉系统已能通过学习抽象特征来识别异常。除了图像级异常检测外，异常定位——即像素级异常检测

也已被积极研究。异常定位不仅提供异常存在与否的信息，还会生成指示异常位置的热力图。需要注意的是，该热力图可作为解释异常成因的起点。

与此同时，异常定位算法无法在学习过程中考虑所有可能的异常值。换言之，它们无法构建一个包含所有异常值的数据集。因此，通过学习正常样本的分布来区分异常样本已成为主流方法。例如，基于无监督学习的方法（如[3,19]）利用了仅用正常特征训练的生成器无法成功重建异常特征这一特性。而基于自监督学习的方法（如[11,21,23]）则通过合成噪声并将其作为异常样本用于学习。近年来，[4,5,16]利用在ImageNet[6]等大型数据集上预训练的CNN构建记忆库，并取得了最先进的性能。这种基于记忆库的方法无需学习目标数据集，即可从预训练CNN中提取足够泛化的特征，并将其存储到记忆库中。最终，通过将输入特征与记忆特征进行匹配，判断输入样本是否异常。

然而，工业图像通常具有与ImageNet不同的分布。因此，预训练的CNN从新的工业图像中提取的特征并不适用。这在异常定位中可能是一个致命问题，因为它需要精确区分正常和异常特征。[16]指出了预训练CNN提取有偏差特征所导致的不匹配问题。该方法仅使用了偏差相对较小的中层特征，但并未从根本上解决不匹配问题。

异常定位的性能取决于记忆库的大小。传统方法会在记忆库中尽可能多地存储目标数据集的正常特征，以适应未拟合的特征，即理解正常特征的分布。因此，记忆库的大小通常与此成比例确定。

目标数据集中。然而，内存库中大量未拟合的特征可能导致异常特征的正态性被高估的风险。此外，大量的内存库会增加推理时间。

为了获得具有判别性的正常特征，我们提出了一种新颖方法，通过将迁移学习应用于预训练的CNN来生成目标导向且偏差减小的特征。首先，我们定义了一种基于软边界回归的新型损失函数，该函数通过搜索最小半径的超球面来密集聚类正常特征。所提出的损失函数利用多个形成耦合超球面的记忆特征，帮助可学习的局部描述符提取判别性特征。其次，为减少推理时间，我们提出了一种可扩展的记忆库。由于该可扩展记忆库独立于目标数据集的大小，它不仅降低了异常特征被高估为正常的风险，还实现了空间复杂度的效率优化。因此，所提方法能够通过为目标数据集提取恰当的目标导向特征，并构建包含核心正常特征的降尺度记忆库，从而有效定位异常。

我们使用MVTec AD基准[1]评估了所提出的方法，这是一个用于视觉检测的流行工业图像数据集。该方法在异常检测性能指标（即图像级AUROC（I-AUROC））上达到了99.5%的性能，并在异常定位性能指标（即像素级AUROC（P-AUROC））上实现了98.5%的SOTA性能。特别值得注意的是，所提出的方法在将内存库[7]中约99.9%的激活减少的同时，提供了优于传统方法的性能。

贡献。本文的贡献总结如下：1) 我们发现预训练CNN中的偏置特征对异常定位存在负面影响，并提出通过目标数据集适配作为解决方案。2) 我们提出一种通过度量学习获取判别性特征的新方法，并通过实验验证该特征能够实现高度精细的异常定位。3) 通过特征适配构建的、与目标数据集规模无关的压缩记忆库，在存储容量大幅降低的情况下仍实现了SOTA性能。

## 2. 相关工作

通常，获取异常样本需要大量成本，且无法考虑所有类型的异常。因此，一种基于记忆库的方法应运而生，该方法通过使用预训练的CNN推断目标数据集来获取正常特征。[4]从特征图中提取正常特征并将其存储在记忆库中。在测试阶段，通过

计算内存库中的正常特征与测试样本中的补丁特征之间的欧几里得距离。[5]通过对特征图每个位置的正态分布进行建模来定义内存库。为了进一步考虑特征间的相关性，该方法采用马哈拉诺比斯距离度量来计算异常分数。[16]仅使用中层特征图来减轻特征偏差，并通过考虑每个正常特征的相邻特征来最大化正常信息。此外，它提出了贪心核心集子采样方法，以减轻内存库的负担并降低时间/空间复杂度。然而，上述方法的共同点是它们使用了基于大型数据集的偏差特征而未进行适配。同时，内存库的大小仍然与目标数据集的大小成正比，且存在无法将内存库调整至任意大小的问题。

## 3. 提出的方法

本文提出了一种称为基于耦合超球面的特征适应（CFA）方法，通过对目标数据集进行迁移学习，以缓解预训练卷积神经网络（CNN）的偏差问题。CFA的补丁描述符通过学习目标数据集中正常样本提取的补丁特征，使其在记忆特征周围具有高密度分布。因此，CFA解决了使用预训练CNN时异常特征被过度估计为正常性的问题。

如图1所示，CFA通过基于预训练CNN（即带有大型数据集的偏置CNN）对目标数据集样本进行推断，获取多尺度特征图。由于CNN各深度采样的特征图具有不同空间分辨率，它们会如文献[5]所述被插值至相同分辨率后进行拼接，从而生成块特征 $\mathcal{F} \in \mathbb{R}^{D \times H \times W}$ 。此处 $H$ 和 $W$ 分别表示最大特征图的高度与宽度， $D$ 代表采样特征图的维度总和。由于 $\mathcal{F}$ 的每个像素位置都具有预定的感受野，块特征 $\mathbf{p}_{t \in \{1, \dots, HW\}} \in \mathbb{R}^D$ 可视为该像素位置的语义信息。接着将 $\mathbf{p}$ 输入块描述符 $\phi(\cdot)$ ： $\mathbb{R}^D \rightarrow \mathbb{R}^{D'}$ 。其中 $\phi(\cdot)$ 是具有可学习参数的辅助网络，可将 $\mathbf{p}_t$ 转换为面向目标的特征 $\phi(\mathbf{p}_t) \in \mathbb{R}^{D'}$ 。此处 $D'$ 表示通过 $\phi(\cdot)$ 嵌入的 $\phi(\mathbf{p}_t)$ 维度。

同时，所有从仅包含正常样本的训练集中获取的初始目标导向特征，都会根据特定的建模程序存储在记忆库 $\mathcal{C}$ 中。在图1中，虚线表示该步骤仅在初始化阶段执行（参见第3.2节）。在训练阶段，CFA基于以记忆特征 $\mathbf{c} \in \mathcal{C}$ 为中心构建的叠加超球面进行对比监督，即所谓的耦合超球面。需要注意的是， $\phi(\mathbf{p}_t)$ 被训练为

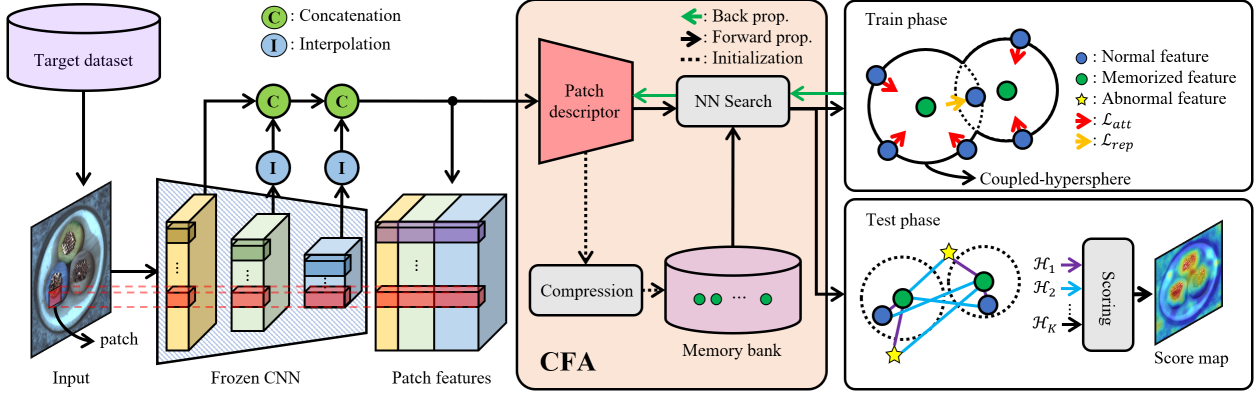


图1. 我们提出的方法（CFA）的整体结构。

在训练阶段密集聚集的特征，即正常特征，对于区分异常特征非常有用。在测试阶段，CFA将测试集任意样本得到的 $\mathbf{p}_t$ 与记忆库中搜索到的最近邻 $\mathbf{c}_t$ 进行匹配，并生成表示异常程度的热力图。最后，通过特定的评分函数（参见第3.3节）从热力图中计算出用于异常定位的得分图。需要注意的是，该得分图显示了异常特征的精细化区域。

本文结构如下：第3.1节定义了 $\phi(\cdot)$ 并说明其训练方法，第3.2节定义了 $\mathcal{C}$ ，最后第3.3节展示了异常分数的计算过程。

### 3.1. 基于耦合超球面的特征适配

本节描述了如何通过基于记忆库的迁移学习来学习附着在预训练CNN上的 $\phi(\cdot)$ ，以实现更精细的目标导向异常定位。先前的研究[18]和[23]通过引入超球面概念来学习目标数据集的分布，但由于未使用记忆库，仍存在无法清晰理解正常特征的问题。因此，我们提出了一种有效学习 $\phi(\cdot)$ 的方法。所提出的方法通过融合基于超球面的损失函数和记忆库，能够解决预训练CNN的偏差问题。具体流程如下：

为了获得能够清晰检测异常特征的特征空间，我们提取了聚类后的正常特征，使得 $\phi(\cdot)$ 具有高密度。首先，通过 $\phi(\mathbf{p}_t)$ 和 $\mathcal{C}$ 的最近邻搜索，找到第 $k$ 个最近邻 $\mathbf{c}_t^k$ 。接着，CFA监督 $\phi(\cdot)$ ，使得 $\mathbf{p}_t$ 被嵌入到接近 $\mathbf{c}_t^k$ 的位置。具体而言， $\phi(\cdot)$ 通过监督 $\mathbf{p}_t$ 将其嵌入到以 $\mathbf{c}_t^k$ 为中心、半径为 $r$ 的超球体内，从而能够在正常特征之间形成高浓度。因此， $\mathcal{L}_{att}$ 通过向 $\phi(\mathbf{p}_t)$ 添加惩罚来吸引，直至

$r$  远离  $\mathbf{c}_t^k$  的描述为

$$\mathcal{L}_{att} = \frac{1}{TK} \sum_{t=1}^T \sum_{k=1}^K \max\{0, \mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k) - r^2\} \quad (1)$$

其中超参数 $K$ 是与 $\phi(\mathbf{p}_t)$ 匹配的最近邻数量，而 $T = h \times w$ 是从单个样本中获得的 $\mathbf{p}$ 的数量。 $\mathcal{D}(\cdot, \cdot)$ 是预定义的距离度量，即本文中的欧几里得距离。公式

(1) 引导 $\phi(\mathbf{p}_t)$ 逐渐逼近以 $\mathbf{c}_t^k$ 为中心创建的超球面。因此，CFA通过优化 $\phi(\cdot)$ 的参数以最小化 $\mathcal{L}_{att}$ ，通过迁移学习实现特征适应。这样一来，如果通过使用 $\mathcal{L}_{att}$ 的特征适应使 $\phi(\mathbf{p})$ 密集聚集，将易于与异常特征区分开来。

然而，同时属于多个超球面的模糊 $\phi(\mathbf{p})$ 仍可能导致异常特征的正态性被高估。为解决这一问题，我们额外引入硬负样本特征进行对比监督，以获得更具判别力的 $\phi(\mathbf{p}_t)$ 。硬负样本特征定义为通过最近邻搜索与 $\mathcal{C}$ 匹配的、 $\mathbf{p}_t$ 的第 $K+j$ 个最近邻 $\mathbf{c}_t^j$ 。因此，我们定义 $\mathcal{L}_{rep}$ 对 $\phi(\cdot)$ 进行对比监督，使得以 $\mathbf{c}_t^j$ 为中心构建的超球面排斥 $\mathbf{p}_t$ ，具体如下：

$$\mathcal{L}_{rep} = \frac{1}{TJ} \sum_{t=1}^T \sum_{j=1}^J \max\{0, r^2 - \mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^j) - \alpha\} \quad (2)$$

其中超参数 $J$ 是用于对比监督的硬负特征总数，超参数 $\alpha$ 用于控制 $\mathcal{L}_{att}$ 和 $\mathcal{L}_{rep}$ 之间的平衡。

因此，CFA通过使用公式(1)和(2)进行迁移学习，共同优化了 $\phi(\cdot)$ 的参数：

$$\mathcal{L}_{CFA} = \mathcal{L}_{att} + \mathcal{L}_{rep} \quad (3)$$

如果 $\mathbf{c}_t^j$ 和 $\mathbf{c}_t^k$ 与 $\mathbf{p}_t$ 匹配的距离比 $r$ 更近，则公式3中的 $\mathcal{L}_{CFA}$ 直接基于此监督 $\mathbf{p}_t$ 。

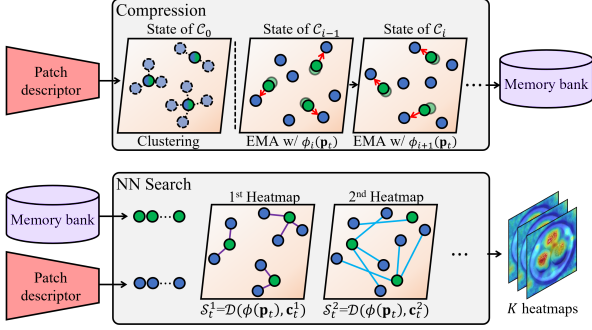


图2. (上) 初始建模记忆库的过程 (下) 通过特征匹配生成热图的过程。

耦合超球面。因此， $\mathbf{p}_t$ 可以通过 $\phi(\cdot)$ 嵌入，使得 $\mathbf{c}_t^k$ 的超球面通过使用 $\mathbf{c}_t^j$ 的对比监督获得更高密度。需要注意的是，我们将通过使用 $\mathcal{L}_{CFA}$ 的迁移学习从补丁描述符中获取目标导向特征的过程命名为“基于耦合超球面的特征适配”。

### 3.2. 记忆库压缩

通过提出的CFA进行迁移学习需要一个记忆库来有效适应目标数据集。然而，如表1所示，先前方法[4, 5, 16]中的建模过程复杂度或记忆库空间往往与目标数据集的大小（即 $\{v^*\}$ ）成比例增加。为了缓解这一现象，本节提出了一种压缩方案来构建高效的记忆库。

记忆库的压缩过程由算法1描述。首先，通过对训练集 $\mathcal{X}$ 中首个正常样本 $\mathbf{x}_0$ 提取的所有 $\phi_0(\mathbf{p}_{t \in \{1, \dots, T\}})$ 应用K-means聚类，构建初始记忆库 $\mathcal{C}_0$ 。在 $\mathcal{C}_0$ 之后更新记忆库的过程如下：推断第 $i$ 个正常样本 $\mathbf{x}_i$ ，并从第 $i-1$ 个记忆库 $\mathcal{C}_{i-1}$ 中搜索最近邻的补丁特征集合 $\mathcal{C}_i^{NN}$ 。接着，通过 $\mathcal{C}_i^{NN}$ 和 $\mathcal{C}_{i-1}$ 的指数移动平均（EMA）计算下一状态 $\mathcal{C}_i$ 的第 $i$ 个记忆库。最终记忆库 $\mathcal{C}$ 是通过训练集所有正常样本重复上述过程 $|\mathcal{X}|$ 次获得的。

图2的上半部分说明了针对目标数据集的每个样本更新记忆特征的过程。遗憾的是，通过 $K$ -NN搜索初始化的 $\mathcal{C}_0$ 并不能整体代表 $\mathcal{X}$ 。然而，若通过指数移动平均法沿 $\phi_i(\mathbf{p}_t)$ 迭代更新 $\mathcal{C}_{i-1}$ ，最终得到的 $\mathcal{C}$ 便能存储代表 $\mathcal{X}$ 的核心正常特征。

由于提出的算法1在每个状态都会更新 $\mathcal{C}$ ，建模过程需要与 $\mathcal{O}(HWD')$ 一样多的空间复杂度。此外， $\mathcal{C}$ 具有与聚类中心数量相同的特征维度 $D'$ 的 $\phi(\mathbf{p})$ ，因此它拥有与

### Algorithm 1 Memory Bank Modeling.

**Require:** Patch descriptor  $\phi$ , dataset  $\mathcal{X}$ , EMA parameter  $\beta$   
**Initialization:**  $\mathcal{C}_0 \leftarrow \text{KMeans}\phi_0(\mathbf{p})$   
**for**  $i \in \{1, \dots, |\mathcal{X}|\}$  **do**  
 $\mathcal{C}_i^{NN} \leftarrow \{\}$   
**for**  $j \in \{1, \dots, |\mathcal{C}|\}$  **do**  
 $Y \leftarrow (\phi_i(\mathbf{p}) \cup \mathcal{C}_i^{NN}) \cap (\mathcal{C}_{i-1}^{NN})^c$   
 $\mathcal{C}_i^{NN} \cup \arg \min_{y \in Y} \|y - \mathcal{C}_{i-1}^j\|_2$   
**end for**  
 $\mathcal{C}_i \leftarrow (1 - \beta) \cdot \mathcal{C}_{i-1} + \beta \cdot \mathcal{C}_i^{NN}$   
**end for**  
 $\mathcal{C} \leftarrow \mathcal{C}_{|\mathcal{X}|}$   
**return**  $\mathcal{C}$

表1. 内存库建模与内存库大小的复杂度估计。

Methods	Modeling	Memory Bank
SPADE	$\mathcal{O}( \mathcal{X} HWD)$	$\mathcal{G} \in \mathbb{R}^{ \mathcal{X}  \times H \times W \times D}$
PaDiM	$\mathcal{O}( \mathcal{X} HWD^2)$	$\mathcal{N}(\mu, \Sigma) \in \mathbb{R}^{H \times W \times D^2}$
PatchCore	$\mathcal{O}( \mathcal{X} HWD')$	$\mathcal{M} \in \mathbb{R}^{ \mathcal{X}  \times \gamma(H \times W) \times D'}$
Ours	$\mathcal{O}(HWD')$	$\mathcal{C} \in \mathbb{R}^{\gamma(H \times W \times D)}$

与 $\mathcal{O}(\gamma(HW)D')$ 一样高的空间复杂度。这里，压缩比 $\gamma$ 表示 $T$ 的比率，即从 $\mathcal{F}$ 获得的 $\mathbf{p}$ 数量与聚类中心数量的比值。因此，如表1所示，所提出方法的 $\mathcal{C}$ 不受 $|\mathcal{X}|$ 的影响。

### 3.3. 评分函数

在3.1节中，使用 $\mathcal{D}(\cdot, \cdot)$ 计算了 $\phi(\mathbf{p}_t)$ 和 $\mathbf{c}_t^k$ 之间的距离。 $\min_k \mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k)$ 表示 $\phi(\mathbf{p}_t)$ 与记忆特征 $\mathcal{C}$ 之间的最小距离，即 $\phi(\mathbf{p}_t)$ 的异常程度。因此，我们可以直接使用 $\mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k)$ 来朴素地定义异常分数，如下所示：

$$S_t = \min_k \mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k) \quad (4)$$

然而，由于正常特征是连续分布的，聚类之间的边界并不清晰。因此，仅凭朴素异常评分难以精确区分异常特征。具体而言，难以确定哪些记忆特征会与 $\phi(\mathbf{p}_t)$ 匹配。在这种情况下，即使 $\phi(\mathbf{p}_t)$ 是正常特征，它也可能位于记忆特征的中间位置，导致距离较大。因此，仅基于距离的朴素异常评分可能会低估正常特征的正常性。为此，我们提出了一种新的评分函数，该函数考虑了 $\phi(\mathbf{p}_t)$ 的不确定性。

$\phi(\mathbf{p}_t)$ 匹配得越清晰，与特定记忆特征的距离相比其他记忆特征就越近。

表2. MVTec AD数据集上异常定位方法的图像/像素级AUROC (%)。

Model		SPADE	Patch SVDD	PaDiM	CutPaste	CFLOW	PatchCore	CFA	CFA++
I-AUROC	Textures	96.6	94.5	95.3	98.4	98.7	99.0	99.6	<b>99.8</b>
	Objects	96.0	90.8	95.3	94.1	98.0	99.1	99.2	<b>99.4</b>
	All	96.2	92.1	95.3	95.5	98.3	99.1	99.3	<b>99.5</b>
P-AUROC	Textures	92.9	93.7	95.3	96.9	<b>98.5</b>	97.5	97.2	97.5
	Objects	97.6	96.7	95.3	97.8	98.7	98.3	98.6	<b>98.9</b>
	All	96.0	95.7	97.5	97.5	<b>98.6</b>	98.2	98.2	98.5

因此，我们使用softmin来衡量最近的 $c$ 与其他 $c$ 相比的接近程度，并将其定义为确定性。这样一来，通过将 $S_t^k$ 与 $\phi(\mathbf{p}_t)$ 的确定性相乘，低估正态性的问题得以解决。公式描述如下：

$$\mathcal{A}_t = \frac{e^{-S_t}}{\sum_{k=1}^K e^{-\mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k)}} \cdot S_t \quad (5)$$

最后，在CFA的测试阶段，异常分数图（即异常定位的最终输出）是从热图中获得的。需要注意的是，这些热图是由原始异常分数生成的，如图2下半部分所示。简而言之，首先生成第 $k$ 个热图 $\mathcal{H}^k = \{\mathcal{D}(\phi(\mathbf{p}_t), \mathbf{c}_t^k) | 1 \leq t \leq T\}$ 并重新排列，使 $\mathcal{H}^k$ 具有空间信息。随后，在所有像素位置计算式(5)，得到CFA的最终输出——异常分数图 $\mathcal{A}$ 。此处，为了输出与输入样本分辨率相同的异常分数图，对 $\mathcal{A}$ 进行了适当插值，并应用了 $\sigma = 4$ 的高斯平滑作为后处理步骤。

总而言之，CFA利用提出的补丁描述符和记忆库，针对目标导向的异常定位进行迁移学习。随后，CFA从任务导向的特征生成热力图，并据此计算精细的异常分数。因此，CFA解决了因预训练CNN的偏置特征导致异常特征正常性被高估的问题。

## 4. 实验

本节展示了各种实验结果，以评估CFA的异常检测与定位性能。所有实验均在MVTec AD基准数据集[1]上进行，该数据集是异常定位领域最著名的数据集。为验证所提方法的鲁棒性，我们还展示了在随机旋转裁剪后的Rd-MVTec AD数据集上的性能表现。评估指标采用受试者工作特征曲线下面积（AUROC），并从异常检测（I-AUROC）和定位（P-AUROC）两个维度评估所提方法的性能。部分实验中，我们使用了面积

在Per-Region-Overlap曲线（P-AUPRO）[2]下，可以更精确地评估异常定位。

### 4.1. 实验设置

本节描述了本文实验所设置的配置。实验中使用的CNN均使用ImageNet进行预训练。为了在预训练的CNN上获取多尺度特征，我们按照[12]的方法，从中间层提取对应于 $\{C_2, C_3, C_4\}$ 的特征图。每个提取的特征图的空间分辨率分别为输入样本的1/4、1/8和1/16。特别地，对于通道维度非常小的EfficientNet，每个尺度使用了多个特征图，这些特征图按通道维度值进行划分。采用1×1 CoordConv层[13]作为补丁描述符，其参数使用He初始化器[9]进行初始化。为优化补丁描述符的参数，使用了AdamW[14]优化器，并应用了amsgrad[15]。学习率设置为1e-3，未使用任何调度器，权重衰减设置为5e-4。批量大小设置为4。补丁描述符训练30个周期，每个子类约需10分钟。CFA的超参数中， $r$ 和 $\alpha$ 分别设置为1e-5和1e-1。每个补丁特征的最近邻数量 $K$ 和 $J$ 均设置为3。实验使用Quadro RTX 5000 GPU和Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz CPU来测量所提方法的吞吐量。

实验中使用的MVTec AD数据集是最大的工业样本数据集，包含5354个样本，其中1725个为测试样本。该数据集分为15个子类，我们对每个类别独立进行迁移学习。在预处理阶段，数据集的每个样本被调整为256×256，并进行中心裁剪至224×224。此外，我们使用RD-MVTec AD数据集进一步考虑未对齐样本，这些样本更难检测异常值。RD-MVTec AD数据集的每个样本在±10°范围内随机旋转。随机旋转后，每个样本被调整为256×256，然后随机裁剪至224×224。

表3. RD-MVTec AD数据集上异常定位方法的图像级AUROC (%) 与像素级AUPRO (%)。

Model		Textures	Objects	All
VAE (ResNet18)	I-AUROC	54.7	65.8	62.1
	P-AUPRO	23.1	30.2	27.8
CFA++ (ResNet18)	I-AUROC	<b>98.6</b>	<b>95.5</b>	<b>96.5</b>
	P-AUPRO	<b>81.1</b>	<b>82.2</b>	<b>81.8</b>
SPADE (WRN50-2)	I-AUROC	84.6	88.2	87.2
	P-AUPRO	75.6	65.8	69.0
PaDiM (WRN50-2)	I-AUROC	92.4	92.1	92.1
	P-AUPRO	77.8	70.8	73.1
CFA++ (WRN50-2)	I-AUROC	<b>99.7</b>	<b>98.3</b>	<b>98.7</b>
	P-AUPRO	<b>82.2</b>	<b>83.7</b>	<b>83.2</b>

4.2. 定量结果

本节研究了CFA的量化性能。表2展示了CFA在MVTec AD数据集上使用ImageNet预训练的WRN50-2时取得的I-AUROC和P-AUROC结果。此处CFA++指代同时使用裁剪图像与仅调整尺寸样本的结果集成情况。所提方法在评估图像级异常检测性能的I-AUROC指标上，对纹理类和物体类均取得了SOTA性能。例如CFA++的I-AUROC比当前保持SOTA性能的PatchCore[16]高出0.4%。在评估像素级异常定位的P-AUROC指标方面，所提方法对物体类同样实现了SOTA性能。但当考虑全部类别时，所提方法的P-AUROC略低于CFLOW[8]。尽管如此，传统方法仅在异常检测或定位单一层面表现突出，而所提方法能同时在两种场景下保证优异性能。值得注意的是，相较于SPADE、PaDiM和PatchCore，所提方法在采用空间复杂度更低的内存库情况下，仍通过特征适配实现了卓越性能。表4展示了各架构性能以更清晰地呈现所提方法的性能提升效果。可见CFA++在多数类别中性能最优。特别值得注意的是，CFA与CFA++的最差性能仍达97.3%，显著高于其他多数技术。这种趋势源于CFA通过特征适配获得了适用于多类别的泛化性能。

表3展示了CFA在RD-MVTec AD数据集上的性能。RD-MVTec AD数据集包含与MVTec AD数据集相同的样本，但未经过对齐处理。

因此，该数据集的性能普遍低于MVTec AD。以表3为例进行比较，SPADE的I-AUROC下降了9%。这意味着SPADE对非规整数据集非常敏感。另一方面，即使在未对齐的样本中，CFA++的I-AUROC仅出现0.8%的轻微性能下降，这表明所提方法能像HVS一样稳健地区分正常特征。此外，与SPADE和PaDiM相比，CFA++的I-AUROC分别高出11.5%和6.6%。在评估精细检测的P-AUPRO指标上也观察到类似趋势——例如CFA++的P-AUPRO分数比PaDiM高出10.1%。

4.3. 消融研究

表5展示了特征适应对异常定位的影响。首先来看仅使用预训练CNN偏置特征的情况。未经适应的预训练CNN虽然从大数据集中获得了丰富特征，但由于偏置特征导致性能低下。这是因为偏置特征使正常特征的正态性被低估，从而对异常定位产生负面影响。此时，当仅使用 $\mathcal{L}_{att}$ 时，CFA将ResNet18的I-AUROC和P-AUROC分数分别提升了14.1%和5.4%。对于WRN50-2，借助 $\mathcal{L}_{att}$ 使I-AUROC和P-AUROC分数分别提升了13.2%和4.3%。这是因为正常特征在记忆特征周围更紧密地聚集。然而仍存在一个问题：由于无法确定它们属于哪个超球面，这些特征缺乏区分性。因此，通过引入 $\mathcal{L}_{rep}$ 来获取区分性特征，有望实现进一步的性能提升。实际上在ResNet18中，I-AUROC和P-AUROC分数分别提高了1.1%和0.3%；在WRN50-2中则分别提升了0.4%和0.2%。通过引导正常特征以更具区分性的方式聚集，异常特征得以更精确地区分。另一方面值得注意的是，ResNet18比WRN50-2表现出更大的性能提升。由于ResNet18使用相对较小的特征维度可能增加模糊性，超球面重叠问题会更为严重。 $\mathcal{L}_{rep}$ 有效解决了这一问题。

表6展示了当通过额外采用特征维度缩减比例 $\gamma_d$ 和补丁缩减比例 $\gamma_c$ 对记忆库进行压缩时的P-AUROC分数。首先需注意，CFA的记忆库大小与目标数据集无关。例如，在包含209个样本的Bottle类别中，CFA被压缩至约 $\frac{1}{|\mathcal{X}^{bottle}|}$ 的大小，即约0.5%。也就是说，每个子类的压缩比例 $\gamma$ 计算公式为 $\frac{\gamma_d \gamma_c}{|\mathcal{X}|}$ 。即使在记忆库从25%压缩至约2%的情况下，CFA的P-AUROC分数仅略微下降了0.08%。

表4. MVTec AD数据集中各类别图像级AUROC (%) 性能比较。红色、蓝色和加粗分别代表第一、第二和第三名

Class	SPADE	Patch SVDD	PaDiM	CutPaste	CFLOW	PatchCore	CFA	CFA++
Bottle	-	98.6	-	98.2	<b>100</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Cable	-	90.3	-	81.2	97.6	<b>99.5</b>	<b>99.8</b>	<b>99.8</b>
Capsule	-	76.7	-	<b>98.2</b>	97.7	<b>98.1</b>	97.3	<b>99.2</b>
Carpet	-	92.9	-	93.9	<b>98.7</b>	<b>98.7</b>	97.3	<b>99.5</b>
Grid	-	94.6	-	<b>100.0</b>	<b>99.6</b>	98.2	99.2	<b>99.9</b>
Hazelnut	-	92.0	-	98.3	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Leather	-	90.9	-	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Metal nut	-	94.0	-	99.9	99.3	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Pill	-	86.1	-	94.9	<b>96.8</b>	96.6	<b>97.9</b>	<b>97.9</b>
Screw	-	81.3	-	88.7	91.9	<b>98.1</b>	<b>97.3</b>	<b>97.3</b>
Tile	-	97.8	-	94.6	<b>99.9</b>	98.7	<b>99.4</b>	<b>100.0</b>
Toothbrush	-	<b>100.0</b>	-	99.4	99.7	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Transistor	-	91.5	-	96.1	95.2	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
Wood	-	96.5	-	99.1	99.1	<b>99.2</b>	<b>99.7</b>	<b>99.7</b>
Zipper	-	97.9	-	<b>99.9</b>	98.5	99.4	<b>99.6</b>	<b>99.6</b>
Average	96.2	92.1	95.3	95.5	98.3	<b>99.1</b>	<b>99.3</b>	<b>99.5</b>

表5. 在MVTec AD数据集上，根据 $\mathcal{L}_{att}$ 和 $\mathcal{L}_{rep}$ 所提方法的图像/像素级AUROC (%)。

Backbone	$\mathcal{L}_{att}$	$\mathcal{L}_{rep}$	I-AUROC	P-AUROC
ResNet18			83.7	92.4
	✓		97.8	97.8
	✓	✓	<b>98.9</b>	<b>98.1</b>
WRN50-2			85.9	94.0
	✓		99.1	98.3
	✓	✓	<b>99.5</b>	<b>98.5</b>

表6. 在MVTec AD数据集上，采用额外内存库压缩的所提方法在像素级别的AUROC (%)。

Backbone	$\gamma_d$	$\gamma_c$	P-AUROC	Throughput
WRN50-2	1	1	<b>98.45</b>	48
	1/2	1/2	98.44	93 (1.9x)
	1/4	1/4	98.44	132 (2.8x)
	1/8	1/8	98.36	<b>172</b> (3.6x)

使用这种轻量级记忆库对吞吐量的提升有积极影响。CFA的吞吐量考虑了通过预训练CNN和补丁描述符的前向传播推理时间。我们可以观察到

表7. 在MVTec AD数据集上，采用不同预训练CNN的所提方法在图像/像素级别的AUROC (%)。

Backbone	Method	I-AUROC	P-AUROC
VGG19	DFR	93.8	<b>95.5</b>
	CFA++	<b>96.2</b>	95.3
EffiNet-B5	PaDiM	97.9	97.5
	CFA++	<b>98.8</b>	<b>98.0</b>
ResNet18	CFLOW	96.8	<b>98.1</b>
	CFA++	<b>98.9</b>	<b>98.1</b>

如果在相同的实验环境中进一步压缩记忆库，吞吐量可提升至3.6倍。例如，观察表6的第3行，即使记忆库的激活量减少约99.9%，CFA性能也几乎未下降，吞吐量反而提升至2.8倍。这是因为记忆库被压缩以仅提取 $\{v^*\}$ 的核心特征，并通过适配使这些特征密集聚集。

表7展示了根据预训练CNN的异常检测和定位性能。此处，针对最流行的VGG19 [20]、EfficientNet-B5 [22]和ResNet18 [10]，将CFA与先前方法[5, 8, 17]进行了比较。



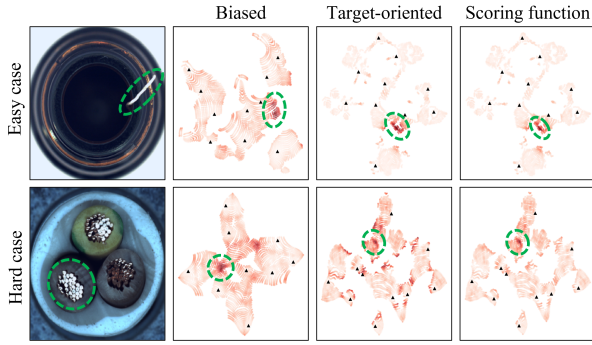


图3. 各补丁特征异常得分的可视化。

CFA在异常定位中得到了广泛应用。与三种预训练的CNN相比，CFA的I-AUROC分数分别高出2.4%、0.9%和2.1%，表现出更优的性能。因此，表7支持了所提出方法的普适性。

#### 4.4. 定性结果

图3展示了根据特征适应和评分函数，每个样本的补丁特征异常得分。此处，红色表示异常得分，虚线圆圈表示异常特征区域，三角形表示记忆特征。在适应前使用偏向大型数据集的特征时，正常特征的正态性被低估，其得分与异常特征相近（见图3第二列）。由于得分边界模糊，难以区分这两种特征，这导致异常特征无法被精确识别的负面影响。

另一方面，当使用特征适应后的目标导向特征时，它们能够很好地聚类。因此，简单案例的正常特征与异常特征被清晰区分（见图3第三列）。然而，仅靠聚类无法精确评估困难案例中不确定的异常特征。所提出的评分函数通过考虑确定性来确定异常分数，因此即使是困难案例中的异常特征也能被精确区分，如图3最后一列所示。因此，所提方法的每一步都有效提升了异常定位性能。

接下来，图4展示了异常定位的结果，这些结果标示出了异常区域。通过CFA获得的异常分数图经过插值处理，以达到输入样本的空间分辨率，并使用 $\{v^*\}_4$ 进行高斯滤波以获得平滑边界。同时，对归一化后的异常分数进行了最小-最大缩放处理。分割结果的阈值是通过计算每个子类所有异常分数的F1得分来确定的。实验结果证明，即使在相当困难的情况下，所提出的方法也能很好地定位异常区域。此外，我们能够

发现所提出的方法在物体和纹理类别中均表现一致。因此，所提出的方法在质量上同样表现出色。

#### 5. 结论

本文指出了主要使用工业图像的异常定位中，预训练CNN引起的偏差问题。为解决此问题，我们提出了基于耦合超球面的特征自适应方法（CFA）以获取目标导向的特征。CFA由一个可与预训练CNN配合使用的可学习块描述符和一个存储记忆特征的记忆库组成。通过迁移学习以及块描述符与预设记忆库关联的特征自适应，CFA成功实现了目标导向的异常定位。在由工业图像构成的最具代表性数据集MVTec AD基准测试中，CFA展现了最先进的性能。随后，通过大量实验从定性/定量角度验证了特征自适应对目标数据集的有效性。

#### 参考文献

- [1] Paul Bergmann, Michael Fauser, David Sattlegger 和 Carsten Steger. Mvtec ad——一个用于无监督异常检测的综合性真实世界数据集。发表于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 第9592–9600页, 2019年。2, 5[2] Paul Bergmann, Michael Fauser, David Sattlegger 和 Carsten Steger. 无信息学生：基于判别性潜在嵌入的师生异常检测。发表于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 第4183–4192页, 2020年。5[3] Paul Bergmann, Sindy Löwe, Michael Fauser, David Sattlegger 和 Carsten Steger. 通过将结构相似性应用于自编码器来改进无监督缺陷分割。 *arXiv preprint arXiv:1807.02011*, 2018年。1[4] Niv Cohen 和 Yedid Hoshen. 基于深度金字塔对应的子图像异常检测。 *arXiv preprint arXiv:2005.02357*, 2020年。1, 2, 4[5] Thomas Defard, Aleksandr Setkov, Angelique Loesch 和 Romaric Audigier. Padim: 一种用于异常检测与定位的补丁分布建模框架。发表于 *International Conference on Pattern Recognition*, 第475–489页。Springer, 2021年。1, 2, 4, 7[6] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li 和 Li Fei-Fei. Imagenet: 一个大规模分层图像数据库。发表于 *2009 IEEE conference on computer vision and pattern recognition*, 第248–255页。Ieee, 2009年。1[7] Piotr Dollár, Mannat Singh 和 Ross Girshick. 快速且准确的模型缩放。发表于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 第924–932页, 2021年。2[8] Denis Gudovskiy, Shun Ishizaka 和 Kazuki Kozuka. Cflow-ad: 实时无监督异常检测与





arXiv:1904.09237, 2019。5 [16] Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, 和 Peter Gehler。迈向工业异常检测的完全召回。arXiv preprint arXiv:2106.08265, 2021。1, 2, 4, 6 [17] Marco Rudolph, Bastian Wandt, 和 Bodo Rosenhahn。相同但又不同：基于归一化流的半监督缺陷检测。于 *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 第1907–1916页, 2021。7 [18] Lukas Ruff, Robert Vandermeulen, Nico Goernitz, Lucas Deecke, Shoaib Ahmed Siddiqui, Alexander Binder, Emmanuel Müller, 和 Marius Kloft。深度单类分类。于 *International conference on machine learning*, 第4393–4402页。PMLR, 2018。3 [19] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, 和 Georg Langs。使用生成对抗网络进行无监督异常检测以指导标记发现。于 *International conference on information processing in medical imaging*, 第146–157页。Springer, 2017。1 [20] Karen Simonyan 和 Andrew Zisserman。用于大规模图像识别的极深度卷积网络。arXiv preprint arXiv:1409.1556, 2014。7 [21] Jouwon Song, Kyeongbo Kong, Ye-In Park, Seong-Gyun Kim, 和 Suk-Ju Kang。Anoseg：使用自监督学习的异常分割网络。arXiv preprint arXiv:2110.03396, 2021。1 [22] Mingxing Tan 和 Quoc Le。EfficientNet：重新思考卷积神经网络的模型缩放。于 *International Conference on Machine Learning*, 第6105–6114页。PMLR, 2019。7 [23] Jihun Yi 和 Sungroh Yoon。Patch SVDD：用于异常检测与分割的补丁级SVDD。于 *Proceedings of the Asian Conference on Computer Vision*, 2020。1, 3