

SimpleNet:一个简单的图像异常检测和定位网络

智康柳 1 号

益铭周 2

盛远 Xu2

中国科学技术

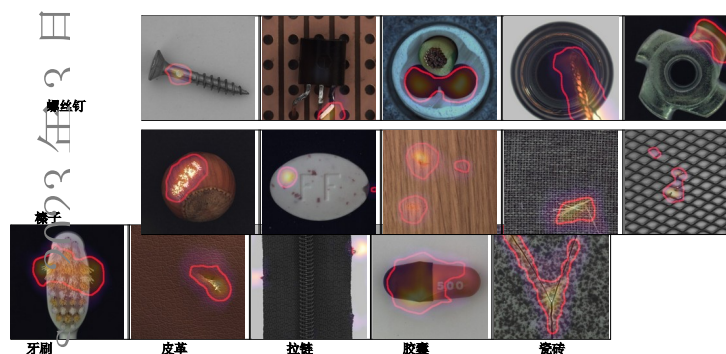
大学自动化系

lzk@mail.ustc.edu.cn

zhouyiming.donal@gmail.com

xys-tc@hotmail.com

zlwang@ustc.edu.cn



图一 MVTec AD 中样本的可视化。产生的异常图叠加在图像上。高异常分数的异常区域用橙色着色。红色边界去除了异常的实际分割图的轮廓。



图二。在 MVTec AD 基准测试中推理速度 (FPS) 与 I-AUROC 的对比。SimpleNet 在准确性和效率上都远远超过了以前的所有方法。

摘要

我们提出了一个简单和应用友好的网络 (称为 SimpleNet) 来检测和定位异常。SimpleNet 包括四个组件: (1) 生成局部特征的预训练特征提取器, (2) 将局部特征传递到目标域的浅层特征适配器, (3) 通过将高斯噪声添加到正常特征来伪造异常特征的简单异常特征生成器, 以及 (4) 将异常特征与正常特征区分开的二元异常鉴别器。在推断过程中, 异常特征生成器将被丢弃。我们的方法基于三个原则。首先, 将预先训练的特征转换成面向目标的特征有助于避免领域偏差。第二, 在特征空间中生成合成异常更有效, 因为缺陷在图像空间中可能没有太多的共性。第三, 简单的鉴别器是非常有效和实用的。尽管简单, SimpleNet 在数量和质量上胜过以前的方法。在

*通讯作者

MVTec AD 基准测试 SimpleNet 的异常检测 AUROC 达到 99.6%, 与下一个性能最佳的模型相比, 误差减少了 55.5%。此外, SimpleNet 比现有方法更快, 在 3080ti GPU 上的帧率高达 77 FPS。此外, SimpleNet 在一类新颖性检测任务中表现出显著的性能改进。代码: <https://github.com/DonaldRR/SimpleNet>。

1. 介绍

图像异常检测和定位任务旨在识别异常图像并定位异常子区域。检测各种感兴趣的异常的技术在工业检测中具有广泛的应用 [3, 6]。在工业场景中, 异常检测和定位尤其困难, 因为异常样本很少, 异常可能从细微的变化 (如细小的划痕) 到大的结构缺陷 (如零件缺失) 不等。来自 MVTec AD 基准测试的一些示例 [3] 以及我们提出的方法的结果如图所示 1。这种情况

禁止被监督的方法接近。

当前的方法以非监督的方式解决这个问题，其中在训练过程中仅使用正常样本。基于重建的方法[10, 21, 31]，基于合成的方法[17, 30]，以及基于嵌入的方法[6, 22, 24]是解决这个问题的三个主要趋势。基于重建的方法，例如[21, 31]假设仅用正常数据训练的深度网络无法准确重建异常区域。逐像素重建误差作为异常定位的异常分数。然而，这种假设并不总是成立的，有时网络可以很好地“概括”，以至于它也可以很好地重建非正常输入，从而导致错误检测[10, 19]。基于综合的方法[17, 30]通过对无异常图像上生成的合成异常进行训练，估计正常和异常之间的判定边界。然而，合成的图像不够逼真。来自合成数据的特征可能偏离正常特征很远，用这样的负样本进行训练可能导致宽松界定的正常特征空间，这意味着不明显的缺陷可能被包括在分布内特征空间中。

最近，基于嵌入的方法[6, 7, 22, 24]实现一流的性能。这些方法使用 ImageNet 预训练的卷积神经网络(CNN)来提取广义正常特征。然后使用统计算法，例如多元高斯分布[6]，正常化流程[24]，以及内存条[22]来实现正态特征分布。通过将输入特征与学习的分布或记忆的特征进行比较来检测异常。然而，工业图像通常具有不同于 ImageNet 的分布。直接使用这些有偏差的特性可能会导致不匹配的问题。此外，统计算法总是遭受高计算复杂度或高存储消耗。

为了缓解上述问题，我们提出了一种新的异常检测和定位网络，称为 SimpleNet。SimpleNet 利用了基于合成和基于嵌入的方式，并做了一些改进。首先，取代直接使用预先训练的特征，我们建议使用特征适配器来产生减少领域偏差的面向目标的特征。第二，代替直接合成图像上的异常，我们提出通过在特征空间中向正常特征添加噪声来生成异常特征。我们认为，通过适当校准的噪声尺度，可以获得一个紧密有界的正常特征空间。第三，我们通过训练简单的鉴别器来简化异常检测过程，这比前述基于嵌入的方法所采用的复杂的统计算法更具有计算效率。具体来说，SimpleNet 利用预先训练的主干进行正常的特征提取，然后是特征适配器

将特征转移到目标域中。然后，通过将高斯噪声添加到适应的正常特征中，简单地产生异常特征。由几层 MLP 组成的简单鉴别器根据这些特征进行训练，以鉴别异常。

SimpleNet 易于训练和应用，具有优异的性能和推理速度。建议的 SimpleNet 基于广泛使用的 WideResnet50 主干网，以 77 fps 的速度运行时，在 MVTec AD 上实现了 99.6 % 的 AUROC，在准确性和效率方面都超过了之前发表的最佳异常检测方法，见图 2。我们进一步将 SimpleNet 引入到一类新颖性检测的任务中，以显示其通用性。这些优势使得 SimpleNet 成为学术研究和工业应用之间的桥梁。代码将公开提供。

2. 相关著作

异常检测和定位方法主要可以分为三类，即基于重构的方法、基于合成的方法和基于嵌入的方法。

基于重建的方法认为异常图像区域不应该能够被正确地重建，因为它们不存在于训练样本中。一些方法[10]利用生成模型，如自动编码器和生成对抗网络[11]来编码和重建正常数据。其他方法[13, 21, 31]作为修补问题的帧异常检测，其中来自图像的补丁被随机掩盖。然后，利用神经网络来预测被擦除的信息。综合结构相似性指数(SSIM)[29]损失函数在训练中应用广泛。异常图被生成作为输入图像和其重建图像之间的逐像素差异。然而，如果异常与正常训练数据共享共同的组成模式(例如，局部边缘)，或者解码器对于很好地解码一些异常编码来说“太强”，则图像中的异常可能被很好地重建[31]。基于合成的方法代表性地，在无异常图像上合成异常。DRE[30]提出了一种网络，该网络以端到端的方式在合成生成的刚好超出分布的模式上被有区别地训练。剪切粘贴[17]提出了一种生成用于异常检测的合成异常的简单策略，该策略剪切图像块并粘贴在大图像的随机位置。CNN 被训练来从正常和扩充的数据分布中区分图像。然而，合成异常的出现与真实异常并不完全相符。实际上，由于缺陷是多种多样且不可预测的，生成一个包含所有异常值的异常集是不可能的。不是合成图像上的异常，而是利用所提出的简单网络，负样本

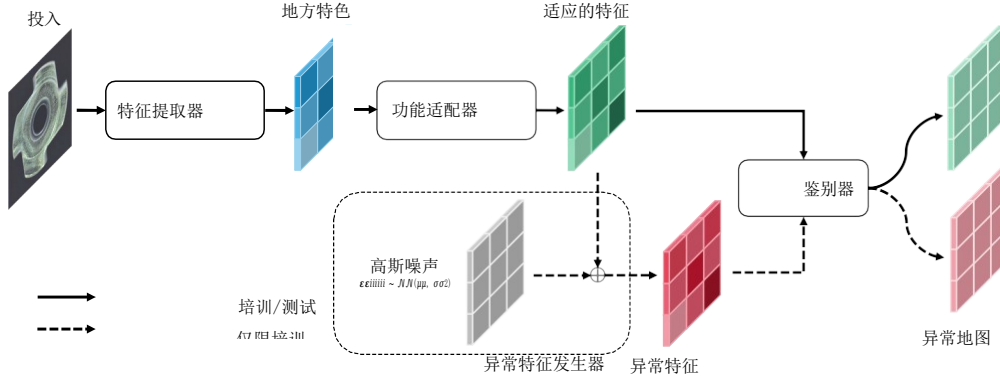


图 3. 提议的简单网络概述。在训练阶段，名义样本被送入预先训练的特征提取器以获得局部特征。然后，利用特征适配器将预训练的特征适配到目标域中。通过将高斯噪声添加到适应的特征来合成异常特征。并且分别用作正样本和负样本来训练最终的鉴别器。推断时移除异常特征生成器。

在特征空间中被合成。

最近，基于嵌入的方法达到了最先进的性能。这些方法将正常特征嵌入到压缩空间中。异常特征远离嵌入空间中的正常簇。典型方法 [6, 7, 22, 24] 利用在 ImageNet 上预先训练的网络进行特征提取。通过预先训练的模型，PaDiM [6] 通过多元高斯分布嵌入提取的异常斑块特征。PatchCore [22] 使用名义补丁特征的最具代表性的记忆库。测试中采用马氏距离或最大特征距离对输入特征进行评分。但是，工业图像的分布通常与 ImageNet 不同。直接使用预先训练的特征可能会导致不匹配的问题。此外，计算协方差的倒数 [6] 或在存储体中搜索最近的邻居 [22] 限制了真实的-

时间性能，尤其是边缘设备。
CS-流量 [24]，CFLow-AD [12]，和 different [23]

建议将正态特征分布转换为

通过标准化流量 (NF) 的高斯分布 [20]。由于标准化流程只能处理全尺寸特征图，即不允许向下采样，并且耦合层 [9] 消耗的内存是普通卷积层的几倍，这些方法消耗内存。蒸馏方法 [4, 7] 训练一个学生网络，使其输出与一个固定的预训练教师网络的输出相匹配。在给定异常查询的情况下，应该检测到学生和教师输出之间的差异。这

由于输入图像要经过老师和学生，计算复杂度加倍。
SimpleNet 克服了上述问题。SimpleNet 使用一个功能适配器来执行传输

在目标数据集上学习以减轻预先训练的 CNN 的偏差。SimpleNet 提出合成 anoma-

而不是直接在图像上。SimpleNet 在推理时遵循单流方式，并且完全由传统的 CNN 块构造，这有利于快速训练、推理和工业应用。

3. 方法

本节将详细介绍提议的 SimpleNet。如图所示 3 SimpleNet 由特征提取器、特征适配器、异常特征生成器和鉴别器组成。异常特征生成器仅在训练期间使用，因此 SimpleNet 在推断时遵循单流方式。这些模块将在下面依次描述。

3.1. 特征提取器

特征提取器获取局部特征，如 [22]。我们将该过程重新表述如下。我们用火车来表示-

将集合和测试集合作为 X_{train} 和 X_{test} 。对于任何图像 X ， $X \in R^{H \times W \times 3}$

$x_i \in R$ 在 X_{train} 网络 X_{test} ，预先训练的网络

ϕ 从不同的层次结构中提取特征，就像通常对类 ResNet 主干所做的那样。由于预先训练的网络偏向于它被训练的数据集，所以只为目标数据集选择级别的子集是合理的。形式上，我们定义 L 子集包括使用的层次索引。来自级别的特征映射

$l \in L$ 表示为 $\phi_l, i \in \phi_l(x_i) \in R^{h_l \times w_l \times c_l}$ 其中 H_l, w_l, c_l 为特征图的高度、宽度和通道尺寸。对于位置 (h, w) 处的入口 $\phi_l, i \in R^{c_l}$ ，其斑块大小为 p 的邻域被定义为

$$N(h, w) = \{(h', w') \mid \begin{aligned} &h' \in [h \lfloor p/2 \rfloor, \dots, h + \lfloor p/2 \rfloor], \\ &w' \in [w \lfloor p/2 \rfloor, \dots, w + \lfloor p/2 \rfloor] \end{aligned}\} \quad (1)$$

聚合 $N(h, w)$ 邻域内的要素
使用聚集函数 fagg(使用自适应平均池-

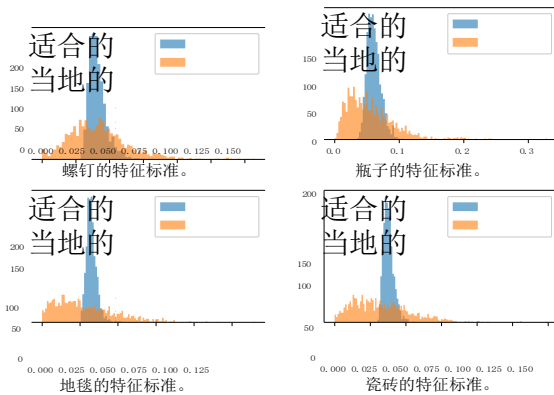


图 4. 沿局部特征和适应特征的每个维度的标准偏差直方图。当用异常特征训练时，适应的特征空间变得更加紧凑。

这里的 ϕ 产生局部特征 $Z1, I$ ，作为

$$\begin{matrix} \text{我, 我} \\ \text{高, 宽} \end{matrix} = \text{法格} \quad | (h', y') \in N(h, w) \quad (2)$$

为了组合来自不同层次的特征 $Z1, I$ ，所有的特征映射被线性地调整到相同的大小 $(H0, W0)$ ，即最大特征映射的大小。简单地按通道连接特征图，得到特征图 $oi \in R^{H0 \times H0 \times C}$ 。该过程定义为

$$oi = \begin{matrix} \text{我} \\ \text{高, 宽} \end{matrix} \text{调整大小} (z) \quad | \quad (h', y') \in N(h, w) \quad (3)$$

我们定义 $oi \in RC$ 作为 oi 在位置 (h, w) 的入口。
我们将上述表达式简化为

$$oi = F\phi(xi) \quad (4)$$

其中 $F\phi$ 是特征提取器。

3.2. 功能适配器

由于工业图像通常具有与主干训练中使用的数据集不同的分布，我们采用特征适配器 $G\theta$ 来传递训练特征

到目标域。功能适配器 $G\theta$ 本地项目

特征 oh, w 为适应特征 qh, w 为

$$\begin{matrix} \text{我} \\ \text{高, 宽} \end{matrix} = G\theta(o_i) \quad (5)$$

特征适配器可以由简单的神经块组成，例如全连接层或多层 perceptron (MLP)。我们通过实验发现，单个全连接层产生良好的性能。

分布估计很难。而 [17, 18, 30] 依靠额外的数据来合成缺陷图像，我们在特征空间的正常样本上添加简单的噪声，声称它优于那些被操纵的方法。

异常特征是通过添加高斯函数产生的
西安噪音对正常的功能齐 $\in RC$ 。形式上，
一个

噪声向量 $\epsilon \in RC$ 被采样，每个条目遵循独立同分布高斯分布 $n(\mu, \sigma^2)$ 。反常现象
特征 qi 融合为

$$\begin{matrix} ih \\ , \\ w \end{matrix} = \begin{matrix} \text{我} \\ \text{高, 宽} \end{matrix} + \epsilon \quad (6)$$

数字 4 说明了异常特征对四类 MVTec AD 的影响。我们可以看到，沿着适应特征的每个维度的标准偏差趋于一致。因此，当区分异常特征和正常特征时，特征空间趋于紧凑
不良特征。

3.4. 鉴别器

鉴别器 $D\psi$ 作为正态性计分器，直接估计每个位置 (h, w) 的正态性。因为负样本是与正常特征一起生成的
 $\{qi | xi \in X_{train}\}$ ，都是训练时喂给判官的。鉴别器期望正常特征的输出为正，而异常特征的输出为负

特色。我们简单地使用一个 2 层多层感知器 (MLP) 结构作为普通分类器，估计 $normality$

$normality$ 为 $D\psi(qh, w) \in R$ 。

3.5. 损失函数和训练

一个简单的截断 L1 损耗的计算公式如下

$$\begin{matrix} \text{我} \\ \text{高, 宽} \end{matrix} = \max(0, th + D\psi(qi)) + \max(0, th - D\psi(qi)) \quad th \quad (7)$$

$th+$ 和 $th-$ 是防止过拟合的截断项。

默认情况下，它们被设置为 0.5 和 -0.5。培训的目标是

$$L = \min_{\theta} \sum_{xi \in X_{train}} \begin{matrix} \text{左侧} \\ \text{左侧} \end{matrix} \quad (8)$$

我们将实验性地评估提议的截断 L1 损失函数与广泛使用的交叉熵损失

3.3. 异常特征发生器

为了训练鉴别器来估计样本为正态的可能性，最简单的方法是负采样

实验部分。算法中给出了训练程序的伪代码 1。

3.6. 推理和评分功能

样本，即缺陷特征，并对其进行优化

正常样本。缺陷的缺乏使得抽样 5。异常分数由鉴别器提供

异常特征生成器在推断时被丢弃。请注意，其余模块可以堆叠成端到端网络。我们将每个 $x_i \in X_{test}$ 输入到前面提到的特征提取器 $F\phi$ 和特征 Adaptor $G\theta$ 依次得到适应的特征 q_i 就像在平等-

$D \psi$
as

$$\frac{1}{w} \int_{\omega} d\psi(q_i)$$

推断期间异常定位的异常图定义为

$$\text{萨}(\mathbf{x}_i) := \{s_i \mid (h, w) \in W_0 \times H_0\}$$

$$(10)$$

然后对萨(x_i)进行插值以获得输入样本的空间分辨率,并对平滑边界进行σ=4的高斯滤波。因为对于任何大小的异常区域都存在最有响应的点,所以将异常图的最大分数作为每个图像的异常检测分数

我

$$\text{伤心}(\mathbf{x}_i) := \max_{(h, w) \in W_0 \times H_0} \text{嘘}, \text{西}$$

$$(11)$$

4. 实验

4.1. 数据集。

我们在MVTec异常检测基准上进行了大部分实验[3],即异常检测与定位领域著名的数据集。MVTec AD包含5个纹理和10个源于制造的对象类别,总共有5354个图像。数据集由用于训练的正常图像和用于测试的具有各种类型缺陷的正常和异常图像组成。它还为有缺陷的测试图像提供像素级注释。典型图像如图1所示。如在[6,22],调整图像大小并居中裁剪至分别为256×256和224×224。没有应用数据扩充。我们遵循一级分类原则

col,也称为冷启动异常检测,其中我们为每个类别在其各自的正常训练样本上训练一个单类分类器。

我们对CI-FAR10[16],其中包含10类50K训练图像和规模为32×32的10K测试图像。在一类新颖性检测的设置下,一类作为正常数据,其他类作为新奇。

4.2. 评估指标。

图像级异常检测性能通过接收器算子曲线下的标准面积来测量,我们将其表示为I-AUROC,使用产生的异常检测分数SAD(等式11)。对于异常定位,异常图SAL(方程10)用于评估逐像素的AUROC(表示为P-AUROC)。根据以前的工作[6,22],我们在MVTec AD上计算用于检测和定位的类平均AUROC和平均AUROC总体类别。比较基线包括AE-SSIM[3],里亚尔[31],DRE[30],剪切粘贴[17],CS-Flow[24],PaDiM[6],RevDist[7]和PatchCore[22]。

算法1 SimpleNet 训练伪代码,Pytorch 类

F:特征提取器

G:功能适配器

N: i.i.d. 高斯噪声# D:鉴别器

预训练初始化(F)随机初始化(G, D)

对于data_loader中的x:

o = F(x) #正常特征 q = G(o) #适应

特征

q_ = q +随机(N) #异常特征

loss = loss_func(D(q), D(q_)).mean()

loss.backward() #反向传播

F = F.detach() #停止渐变

更新(G, D) # Adam

#损失函数

def loss_func(s, s_): th_ =

-th = 0.5

return max(0, th-s) + max(0, th+s_)

4.3. 实施细节

本节描述了本文中实验的配置实现细节。实验中使用的骨干都用ImageNet进行了预训练[8]。主干的第二和第三中间层例如

等式中的 $l' \in [2, 3]$ 在特征提取器中使用,如[22]当主干是ResNet式的architec-

没错。默认情况下,我们的实现使用WideResnet50作为主干,来自特征提取器的特征维数设置为1536。后面的特征适配器本质上是没有任何偏差的完全连接的层。适配器中FC层的输入和输出特征的尺寸是相同的。异常特征生成器增加了id。

高斯噪声N(0,σ²)到正常特征的每个条目。

σ默认设置为0.015。后续鉴别器

由一个线性图层、一个批量标准化图层、一个泄漏relu(斜率为0.2)和一个线性图层组成。在等式中,th+和th均设为0.57。使用Adam优化器,将特征适配器和鉴别器的学习率分别设置为0.0001和0.0002,权重衰减为0.00001。每个数据集的训练时期设置为160,批次大小为4。

4.4. MVTec AD上的异常检测

MVTec AD异常检测结果如表所示1。图像级异常得分由异常图的最大得分给出,如等式所示11。SimpleNet在15门课程中有9门取得了最高分。对于纹理和对象,SimpleNet分别达到I-AUROC的99.8%和99.5%的新SOTA。SimpleNet实现了显著更高的平均图像异常检测性能,即99.6%的I-AUROC得分。清注意

表 1。SimpleNet 与 MVTec AD 上最先进作品的比较。图像式 AUROC (I-AUROC)和像素式 AUROC (P-AUROC)在每个条目中显示为 I-AUROC%/P-AUROC%。CS-Flow 的 P-AUROC 未记录在 [24]

类型	基于重构的		基于合成的		基于嵌入的				我们的
模型	AE-SSIM	利雅得	DRM	剪切粘贴	CS 流	帕迪姆	RevDist	补丁核心	简单网
地毯	87/64.7	84.2/96.3	97.0/95.5	93.9/98.3	100/-	99.8/99.1	98.9/98.9	98.7/99.0	99.7/98.2
格子	94/84.9	99.6/98.8	99.9/99.7	100/97.5	99.0/-	96.7/97.3	100/99.3	98.2/98.7	99.7/98.8
皮革	78/56.1	100/99.4	100/98.6	100/99.5	100/-	100/99.2	100/99.4	100/99.3	100/99.2
瓷砖	59/17.5	98.7/89.1	99.6/99.2	94.6/90.5	100/-	98.1/94.1	99.3/95.6	98.7/95.6	99.8/97.0
木	73/60.3	93.0/85.8	99.1/96.4	99.1/95.5	100/-	99.2/94.9	99.2/95.3	99.2/95.0	100/94.5
平均值。文字。	78/56.7	95.1/93.9	99.1/97.9	97.5/96.3	99.8/-	95.5/96.9	99.5/97.7	99.0/97.5	99.8/97.5
瓶子	93/83.4	99.9/98.4	99.2/99.1	98.2/97.6	99.8/-	99.1/98.3	100/98.7	100/98.6	100/98.0
电缆	82/47.8	81.9/84.2	91.8/94.7	81.2/90.0	99.1/-	97.1/96.7	95.0/97.4	99.5/98.4	99.9/97.6
胶囊	94/86.0	88.4/92.8	98.5/94.3	98.2/97.4	97.1/-	87.5/98.5	96.3/98.7	98.1/98.8	97.7/98.9
黑泽尔胡特	97/91.6	83.3/96.1	100/99.7	98.3/97.3	99.6/-	99.4/98.2	99.9/98.9	100/98.7	100/97.9
金属螺母	89/60.3	88.5/92.5	98.7/99.5	99.9/93.1	99.1/-	96.2/97.2	100/97.3	100/98.4	100/98.8
药丸	91/83.0	83.8/95.7	98.9/97.6	94.9/95.7	98.6/-	90.1/95.7	96.6/98.2	96.6/97.4	99.0/98.6
螺丝钉	96/88.7	84.5/98.8	93.9/97.6	88.7/96.7	97.6/-	97.5/98.5	97.0/99.6	98.1/99.4	98.2/99.3
牙刷	92/78.4	100/98.9	100/98.1	99.4/98.1	91.9/-	100/98.8	99.5/99.1	100/98.7	99.7/98.5
晶体管	90/72.5	90.9/87.7	93.1/90.9	96.1/93.0	99.3/-	94.4/97.5	96.7/92.5	100/96.3	100/97.6
拉链	88/66.5	98.1/97.8	100/98.8	99.9/99.3	99.7/-	98.6/98.5	98.5/98.2	99.4/98.8	99.9/98.9
平均值。Obj。	91/75.8	89.9/94.3	97.4/97.0	95.5/95.8	98.2/-	96.0/97.8	98/97.9	99.2/98.4	99.5/98.4
平均的	87/69.4	91.7/94.2	98.0/97.3	96.1/96.0	98.7/-	95.8/97.5	98.5/97.8	99.1/98.1	99.6/98.1

表二。在不同的 WideResNet50 层次级别组合下，在 MVTec AD 上的性能。

级别 1	级别 2	级别 3	I-AUROC%	P-AUROC%
✓			93.0	94.2
	✓		98.4	96.7
		✓	99.2	97.5
✓	✓		96.7	96.7
	✓	✓	99.6	98.1
✓	✓	✓	99.1	98.1

也就是说，PatchCore 减少了 0.9% 的误差 [22] (下一个最好的竞争对手，在相同的宽度下 50-back-bone) 到 0.4% 意味着误差减少了 55.5%。在工业检测环境中，这是一个相关的重大减少。

4.5. MVTec AD 异常定位

异常定位性能由像素级 AUROC 测量，我们记为 P-AUROC。表中列出了与最先进方法的比较 1。SimpleNet 在 MVTec AD 上实现了 98.1% P-AUROC 的最佳异常检测性能，并且针对对象实现了 98.4% P-AUROC 的新 SOTA。SimpleNet 在 15 个类别中有 4 个类别的得分最高。我们在图中给出了异常定位的代表性样本 8。

4.6. 推理时间

除了检测和定位性能，推理时间是工业模型部署最

重要的考虑因素。与最先进的方法在推理时间上的比较如图所示 2。所有方法都是在相同的硬件上测量的，包括-

使用 Nvidia GeForce GTX 3080ti GPU 和英特尔至强处理器 E5-2680v3@2.5GHZ。它清楚地表明，我们的方法实现了最好的性能以及最快的速度。SimpleNet 比 PatchCore 快近 8 倍[22]。

4.7. 消融研究

邻域大小和等级。我们研究了方程中邻域大小 p 的影响 1。结果如图所示 6 显示了异常预测的局部性和整体性之间的明显的最佳性，因此激发了邻域大小 $p = 3$ 。我们设计了一组实验来测试层次子集 L 对模型性能的影响，结果如表所示 2。我们用 1-3 对前三个 WideResNet50 块进行索引。一样可以可见，来自层级 3 的特征已经可以实现一流的性能，但受益于额外的高层次 2。我们选择 $2 + 3$ 作为默认设置。

适配器配置。适配器提供了一个传输-

预训练特征上的信息(投影)。我们的默认功能适配器是无偏置的单一 FC 层，具有相同的输入和输出通道。表中显示了不同特征适配器的比较 3，第一行“我们的”实施遵循与表中相同的配置 1。“Ours-complex-FA”用非线性特征适配器(即具有非线性的 1 层 MLPs)代替简单特征适配器。行“Ours-w/o-FA”删除特性适配器。结果表明，单个 FC 层可产生最佳性能。直观地，特征适配器找到一个投影，使得伪造的异常特征和投影的预训练特征容易被分离，这意味着鉴别器的简单解决方案。使用特征适配器有助于网络连接的现象也表明了这一点

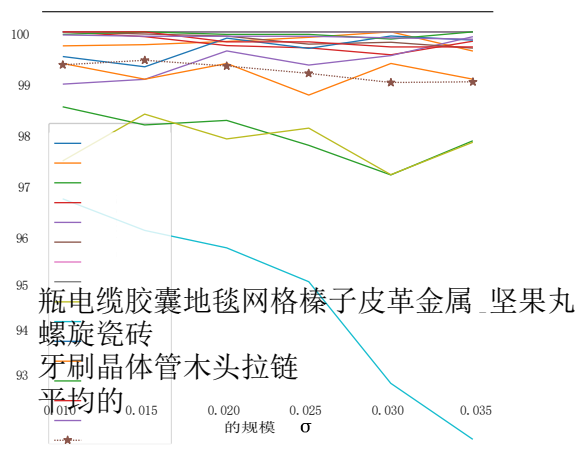


图 5。具有不同 σ 值的各类 MVtec AD 数据集的 I-AUROC%和 P-AUROC%。(彩色观看效果最佳。)

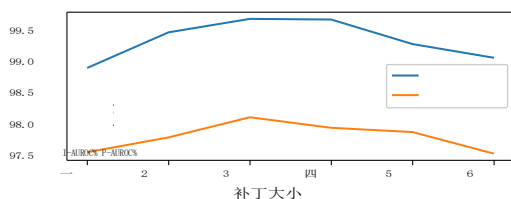
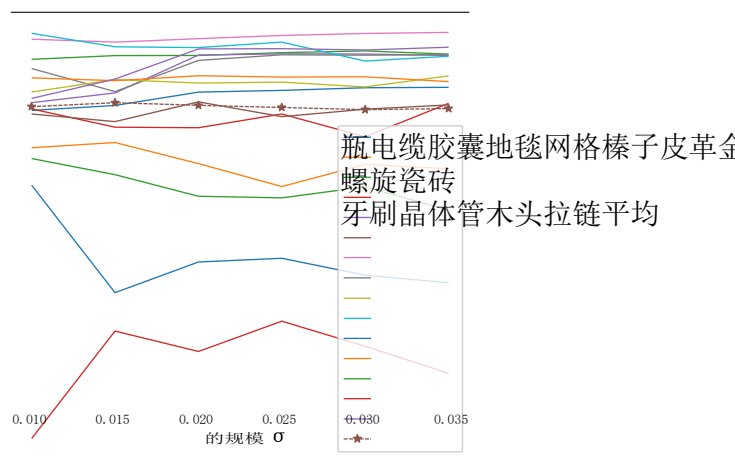


图 6。MVtec AD 上不同补丁大小的性能。

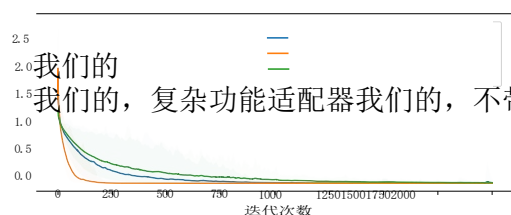


图 7。训练中损失的可视化。绘制的线条显示了 MVtec AD 数据集中所有类别的平均损失。透明色表示损耗波动范围。

边缘快速(图 7)。我们观察到复杂功能适配器的性能显著下降。一个可能的原因是复杂的适配器可能导致过度拟合，降低了测试中各种缺陷的泛化能力。图 4 比较特征适配器前后沿特征每个维度的标准差直方图。我们可以看到，当用异常特征训练时，适应的特征空间变得紧凑。

噪音等级。异常特征发生器中的噪声比例控制合成的异常特征与正常特征的距离。具体来说，高 σ 导致异常特征保持高欧几里得距离

表 3。不同特征适配器的比较。“我们的”实施遵循表中的相同配置 1。“Ours-complex-FA”用非线性特征适配器代替了简单特征适配器。“Ours-w/o-FA”丢弃了特性适配器，相当于使用了一个身份全连接层。“Ours-CE”使用交叉熵损失。显示了 MVtec AD 的 I-AUROC%和 P-AUROC%。

模型	I-AUROC%	P-AUROC%
我们的	99.6	98.1
我们的综合体	98.3	97.2
我们的-w/o-FA	99.2	97.9
我们的-CE	99.4	97.8

表 4。不同主干网下 MVtec 广告的表现。

模型	I-AUROC%	P-AUROC%
ResNet18	98.3	95.7
ResNet50	99.6	98.0
ResNet101	99.2	97.6
宽带网 50	99.6	98.1

表 5。在 CIFAR-10 数据集上的一类新颖性检测 I-AUROC (%) 结果。

方法	电平移动放大器 (Level Shift Amplifier)	DSVDD	奥克甘	HRN	DAAD
奥罗克	64.1	64.8	65.6	71.3	75.3
方法	DisAug CLR	IGD	北马其顿	RevDist	简单网
奥罗克	80.0	83.68	84.5	86.5	86.5

朝向正常特征的距离。在大 σ 上训练将导致宽松的判定界限，导致高假阴性。相反，如果 σ 很小，训练过程将变得不稳定，并且鉴别器不能很好地推广到正常特征。数字 5 详细说明了 σ 对 MVtec AD 中每个类的影响。可以看出， $\sigma = 0.015$ 达到平衡并产生最佳性能。

损失函数。我们比较了提出的损失函数

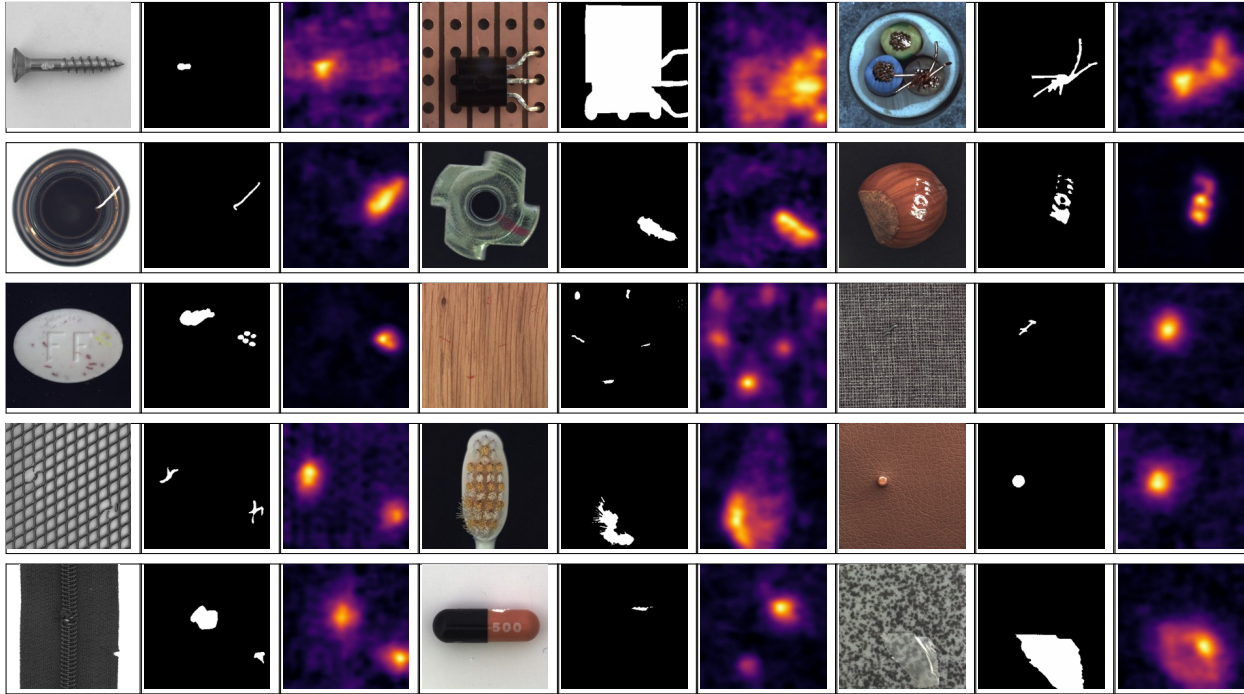


图8。定性结果，其中显示了MVTec AD中每个类别的采样图像、地面实况和异常图。

在截面中 3.5 利用广泛使用的交叉熵损失 (如表中 “Ours-CE” 行所示 3). 我们发现 0.2% I-AUROC 和 0.3% P-AUROC 对交叉熵损失的改善。

对骨干的依赖。我们用不同的主干测试了 SimpleNet，结果如表所示 4。我们发现，在选择不同的主链时，结果大多是稳定的。WideResNet50 的选择与 PaDiM [6] 和 PatchCore [22].

定性结果数字 8 显示异常区域的异常定位结果。通过计算每个子类的所有异常分数的 F1 分数来获得分割结果的阈值。实验结果证明, 该方法即使在非常困难的情况下也能很好地定位异常区域。此外, 我们可以发现, 所提出的方法在对象类和纹理类都具有一致的性能。

4.8. 一类新颖性检测

为了评估所提出的 SimpleNet 的通用性，我们在 CIFAR-10 [16]. 以下[19]，我们使用来自单个类别的样本训练模型，并检测来自其他类别的新样本。我们分别为每个类训练相应的模型。注意，新颖性分数被定义为相似性图中的最大分数。桌子 5 报告我们的方法和其他方法的 I- AUROC 分数。为了公平比较，所有方法都在 ImageNet 上进行了预训练。

基准包括 VAE [2], LSA [1], DSVDD [25], OCGAN [19], HRN [15], 阿诺根 [27], DAAD [14], MKD [26], 禁用 CLR [28], IGD [5] 和 RevDist [7]. 我们的方法优于这些比较方法。请注意, IGD [5] 和 DisAug CLR [28] 分别达到 91.25% 和 92.4%。

5. 结论

在本文中，我们提出了一种简单而有效的方法，称为 SimpleNet，用于无监督的异常检测和定位。SimpleNet 由几个简单的神经网络模块组成，易于训练和应用工业场景。虽然简单，但在 MVtec AD 基准测试中，SimpleNet 实现了最高的性能和最快的推理速度。SimpleNet 提供了一个新的视角，在异常检测和定位的学术研究和工业应用之间架起了一座桥梁。

感谢

本研究得到了国家自然科学基金项目 62176246 和 61836008 的资助。本工作也得到了安徽省自然科学基金 2208085UD17 和中央高校基础研究基金 (WK3490000006) 的资助。

参考

- [1] 大卫·阿巴提、安杰洛·波雷洛、西蒙·卡尔德拉和丽塔·库奇亚拉。新颖性检测的潜在空间自回归。《IEEE/CVF 计算机视觉和模式识别会议论文集》，第 481–490 页，2019 年。**8**
- [2] 金元安和宋祖朝。基于变分自动编码器的异常检测。*IE 专题讲座*, 2(1):1–18, 2015。**8**
- [3] 保罗·博格曼, 迈克尔·福瑟, 大卫·萨特格尔和卡斯滕·斯特格尔。mv tec ad-用于无监督异常检测的综合真实数据集。IEEE/CVF 计算机视觉和模式识别会议论文集, 第 9592–9600 页, 2019 年。**1, 5**
- [4] 保罗·博格曼, 迈克尔·福瑟, 大卫·萨特格尔和卡斯滕·斯特格尔。不知情的学生:带有区别性潜在嵌入的师生异常检测。IEEE/CVF 计算机视觉和模式识别会议论文集, 第 4183–4192 页, 2020 年。**3**
- [5] 、余天、庞冠松和古斯塔沃·卡内罗。基于插值高斯描述子的深度一类分类。《AAAI 人工智能会议论文集》，第 383–392 页，2022 年。**8**
- [6] 托马斯·德法、亚历山大·塞特科夫、安琪丽可·罗施和罗马里奇·奥迪吉耶。Padim:一个用于异常检测和定位的补丁分布建模框架。国际模式识别会议, 475–489 页。斯普林格, 2021。**1, 2, 3, 5, 8**
- [7] 邱含邓和李星宇。基于单类嵌入蒸馏的异常检测。IEEE/CVF 计算机视觉和模式识别会议论文集, 2022 年。**2, 3, 5, 8**
- [8] 邓佳, 魏东, 理查德·索彻, 李·, 和。Imagenet:一个大規模分层图像数据库。2009 年 IEEE 计算机视觉和模式识别会议, 第 248–255 页。Ieee, 2009 年。**5**
- [9] 劳伦特·丁, 雅沙·索尔·迪克斯坦和萨米·本·吉奥。使用真实 nvp 进行密度估计。arXiv 预印本 arXiv:1605.08803, 2016。**3**
- [10] 董功、刘凌桥、Vuong Le、Budhadyta Saha、Moussa Reda Mansour、Svetha Venkatesh 和 Anton van den Hengel。记忆常态以检测异常:用于无监督异常检测的记忆增强深度自动编码器。IEEE/CVF 国际计算机视觉会议论文集, 第 1705–1714 页, 2019 年。**2**
- [11] 伊恩·古德菲勒、让·普吉-阿巴迪、迈赫迪·米尔扎、徐炳、戴维·沃德-法利、谢尔吉尔·奥泽尔、亚伦·库维尔和约舒阿·本吉奥。生成对抗网络。神经网络信息处理系统进展, 27, 2014。**2**
- [12] 丹尼斯·古多夫斯基、石坂顺和友川·小冢。通过条件规范化流定位的实时无监督异常检测。IEEE/CVF 计算机视觉应用冬季会议论文集, 第 98–107 页, 2022。**3**
- [13] 马蒂亚斯·哈塞尔曼、迪特尔·P·格鲁伯和乌尔·塔巴拜。基于深度学习的图像组合异常检测完成。2018 年第 17 届 IEEE 机器学习与应用国际会议 (ICMLA), 第 1237–1242 页。IEEE, 2018。**2**
- [14] 侯、钟巧勇、蒲石良、。分合:学习无监督异常检测的分块记忆。IEEE/CVF 计算机视觉国际会议论文集, 第 8791–8800 页, 2021。**8**
- [15] 胡、、和刘兵。Hrn:一个班级学习的整体方法。神经网络信息处理系统进展, 33:19111–19124, 2020。**8**
- [16] 一个克里日夫斯基。从微小图像中学习多层特征。特龙大学硕士论文, 2009 年。**5, 8**
- [17] 李春良、Kihyuk Sohn、Jinsung Yoon 和 Tomas Pfister。Cutpaste:用于异常检测和定位的自我监督学习。IEEE/CVF 计算机视觉和模式识别会议论文集, 第 9664–9674 页, 2021。**2, 4, 5**
- [18] 菲利普·里兹纳斯基、卢卡斯·拉夫、罗伯特·A·范德穆伦、比利·乔·弗兰克斯、马里乌斯·克洛夫特和克劳斯·罗伯特·穆勒。可解释的深一类分类。2020 年学习表征国际会议。**4**
- [19] Pramuditha Perera、Ramesh Nallapati 和 Bing Xiang。Oc- gan:使用具有受限潜在表示的 gan 进行一类新奇检测。《IEEE/CVF 计算机视觉和模式识别会议论文集》，第 2898–2906 页，2019 年。**2, 8**
- [20] 达尼洛·雷森德和沙基尔·穆罕默德。归一化流的变分推断。国际机器学习会议, 第 1530–1538 页。PMLR, 2015 年。**3**
- [21] nicolae-ca˘ta˘lin·里斯特亚、尼卢·马丹、拉杜·图多尔·约内斯库、卡迈勒·纳斯罗拉希、法哈德·沙赫巴兹·汗、托马斯·b·莫斯-隆德和穆巴拉克·沙阿。用于异常检测的自监督预测卷积注意力块。IEEE/CVF 计算机视觉和模式识别会议论文集, 第 13576–13586 页, 2022。**2**
- [22] Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox 和 Peter Gehler。工业异常检测中的总体召回。IEEE/CVF 计算机视觉和模式识别会议论文集, 第 14318–14328 页, 2022。**2, 3, 5, 6, 8**
- [23] 马可·鲁道夫, 巴斯蒂安·万特和博多·罗森哈恩。相同但不同的是:使用规范化流程的半监督缺陷检测。IEEE/CVF 计算机视觉应用冬季会议论文集, 第 1907–1916 页, 2021。**3**
- [24] 马可·鲁道夫, 汤姆·魏尔宾, 博多·罗森哈恩和巴斯·田·万特。基于图像的缺陷检测的全卷积交叉尺度流。IEEE/CVF 计算机视觉应用冬季会议论文集, 第 1088–1097 页, 2022。**2, 3, 5, 6**
- [25] Lukas Ruff、Robert Vandermeulen、Nico Goernitz、Lucas Deecke、Shoaib Ahmed Siddiqui、Alexander Binder、Emmanuel Müller 和 Marius Kloft。深度一级分类。在机器学习国际会议上, 第 4393–4402 页。PMLR, 2018。**8**

- [26] 穆罕默德·礼萨·萨利希、尼乌莎·萨德贾迪、索罗什·巴塞里扎德、穆罕默德·罗班和哈米德·拉比耶。用于异常检测的多分辨率知识提取。IEEE/CVF 计算机视觉和模式识别会议论文集，第 14902 - 14912 页，2021。[8](#)
- [27] 托马斯·施莱格、菲利普·塞波·克、塞巴斯蒂安·瓦尔德斯坦、乌拉·施密特·埃尔富思和格奥尔格·朗斯。利用生成对抗网络引导标记发现的无监督异常检测。医学成像信息处理国际会议，第 146-157 页。斯普林格，2017。[8](#)
- [28] Kihyuk Sohn, Jinsung Yoon, Minho Jin 和 Tomas Pfister。学习和评估深一类分类的表示。2021 年国际学习表征会议。[8](#)
- [29] 纣王、艾伦·博维克、哈米德·谢赫和埃罗·P·西-蒙塞利图像质量评估:从误差可见性到结构相似性。IEEE 图像处理汇刊, 13(4):600 - 612, 2004。[2](#)
- [30] Vitjan Zavrtanik、Matej Kristan 和 Danijel skocaj。用于表面异常检测的区别训练重建嵌入。IEEE/CVF 计算机视觉国际会议论文集，第 8330 - 8339 页，2021。[2, 4, 5](#)
- [31] Vitjan Zavrtanik、Matej Kristan 和 Danijel skocaj。视觉异常检测中的修复重建。模式识别, 112:107706, 2021。[2, 5](#)