

DOI: 10.13382/j.jemi.B2306646

融合自注意力与缺陷凸显的缺陷检测方法*

乐华钢 王文武 朱磊 朱思宇

(武汉科技大学信息科学与工程学院 武汉 430081)

摘要:针对无监督缺陷检测中重建网络在抑制异常重建的同时无法保留正常区域细节信息的问题,提出了一种融合自注意力与缺陷凸显的缺陷检测方法。首先,在重建网络中引入离散小波变换(DWT)进行下采样,并使用离散小波逆变换(IDWT)进行上采样。相较于传统重建网络,这种方法能减少细节信息的丢失,并对特征进行频率分解。同时,在跳连接中加入自注意力模块对特征重新编码,使其重点关注正常区域的细节信息。此外,设计了一个缺陷区域凸显模块,利用正常样本特征构建特征库,将从测试图像提取的特征与特征库中特征对比得到异常图,将异常图与重建差值图相结合来改善缺陷定位结果。在工业缺陷检测数据集MVTec AD上进行测试,在图像级AUROC上取得了99.3%的结果,同时在像素级AUROC上取得了98.3%的结果,在无监督缺陷检测中具有较高的检测精度和鲁棒性。

关键词:缺陷检测;重建网络;自注意力;缺陷凸显

中图分类号: TP391.41; TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.40

Defect detection method integrating self-attention and highlighting of defects

Yue Huagang Wang Wenwu Zhu Lei Zhu Siyu

(School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China)

Abstract: To address the issue of reconstruction networks in unsupervised defect detection failing to preserve detailed information of normal regions while simultaneously suppressing abnormal reconstructions, a defect detection method that combines self-attention and defect highlighting is proposed. First, the discrete wavelet transform (DWT) is introduced in the reconstruction network for downsampling, and the inverse discrete wavelet transform (IDWT) is used for upsampling. Compared to traditional reconstruction networks, this method reduces the loss of detail information and performs frequency decomposition on features. Then self-attention modules are added into the skip connections to re-encode the features, enabling the features to focus on the details of the normal region. Additionally, a defect region highlighting module is designed, which utilizes features from normal samples to construct a feature library. By comparing the features extracted from the test image with the features in the library, an abnormal map is obtained. Finally, the abnormal map is combined with the reconstruction residual map to improve the results of defect localization. The proposed method is tested on the industrial defect detection dataset MVTec AD and achieved 99.3% area under the receiver operating characteristic curve (AUROC) at the image level and 98.3% at the pixel level, demonstrating high detection accuracy and robustness in unsupervised defect detection.

Keywords: defect detection; reconstruction network; self-attention; defect highlighting

0 引言

随着深度学习的发展,各种机器视觉检测方法已广

泛应用于工业制造中^[1-2]。无监督方法^[3]相较于有监督方法^[4-5],由于无需标注样本且具有更强的适应性而备受关注。其中,以卷积神经网络构成的重建网络发挥着重要作用。重建网络仅在正常样本下进行训练,测试时,正

常样本能够得到完美重建,而异常样本由于未被训练过则会产生较大的重建误差。自动编码器是一种常见的重建网络^[6],它将图像编码成抽象的深层特征,再通过解码器将其重新映射到图像空间。深层特征有利于将异常样本重建为正常,但它所包含正常区域信息也相应地减少,从而影响了对正常区域的精确重建。

为了对正常区域的细节信息进行补充,一些方法^[7]通过增加跳连接来增强自编码器的重构能力。但这种 U-net 结构的网络在只通过正常图像训练时,很容易陷入过拟合,导致缺陷区域也被很好地重建。为了克服这一问题,一些方法^[8-9]使用额外内存空间构建记忆模块,期望在训练时,存储正常样本的特征,在重建时通过检索编码的特征与缓存的内存项之间的相关性权重,来重构编码的特征。然而现有的缓存策略不能保证使用最佳缓存内存项来重建异常区域,并且记忆模块是以像素级的特征作为记忆项,不能很好地考虑到上下文信息。

大部分缺陷在视觉上表现为与周围不同的区域,属于高频特征,与正常区域有不同的频率分布。LIANG 等^[10]采用了高斯差分的方法将输入图像分解为不同的频率成分,并将它们分别进行重建再相加。Wu 等^[11]在缺陷检测领域引入了 DWT 和一种基于 Gumbel softmax 的记忆模块。他们使用 DWT 代替池化层进行下采样并对特征进行频率分解,在解码器中使用 IDWT 代替上采样。小波相较于高斯差分的方式能够同时提取频率和空间信息,并且具有正交性,有利于在重建任务中恢复细节信息。而 Gumbel-softmax 使用 Gumbel 分布和 softmax 函数来近似生成离散分布的采样,通过引入噪声项来增加采样的随机性。但噪声会使得采样结果在某些情况下更加随机或不准确,从而导致采样分布与真实分布的不一致。这种不一致会影响记忆模块的性能,使其不能选择最合适的特征来进行重构。

近年来,基于特征嵌入的方法由于良好的性能越来越受关注。这些方法^[12-15]通过特征提取网络对样本进行特征提取,再比较测试样本和正常样本的特征差异来得到异常图。虽然这种异常图比较粗糙,但通常能大面积地凸显出异常区域。为确保网络具备良好的特征提取能力,这些方法需要在 ImageNet 数据集上进行预训练。由于训练代价较大,所以它们通常采用已经训练好的通用主干作为特征提取网络,但这样会使主干网络的选择受到限制。

针对上述问题,本文设计了一个结合小波变换和自注意力模块的重建网络。自注意力机制可以考虑全局的上下文信息,通过赋予正常区域高权重的方式使网络重点关注正常样本的特征。而小波能代替下采样并对特征进行频率分解,将需要保留的高频细节信息和低频上下文语义信息分开,同时能在一定程度将高频的缺陷特征

与正常特征区分开。为了进一步凸显缺陷区域的重建差值,并减小不正常区域的重建误差对重建差值图带来的影响,本文还设计了一个用于改善重建网络缺陷定位结果的缺陷区域凸显模块,该模块直接利用编码器提取到的特征信息,对最后重建差值图做进一步优化,来提高像素级缺陷检测结果。

1 本文方法

1.1 整体框架

本文方法整体结构如图 1 所示,分为上下两部分。上半部分为重建网络,输入图像先经过编码器中的 DWTblock 依次分解为高频分量 H 和低频分量 L 。由于图像的低频部分包含更多的语义信息,而高频部分包含更多的细节和纹理信息,所以高频分量被送入注意力模块进行重新编码,以保留正常区域的高频信息,而低频部分则继续下采样以获得更抽象的表示。编码器中最底层的低频分量和高频分量在经过注意力模块后,通过解码器中的 IDWTblock 进行融合,得到新的低频分量,继续与后面的高频分量融合,直到得到最后的重建结果。将重建结果与输入图像求差值得到重建差值图,通过差值图能求出重建网络的图像级和像素级的缺陷检测结果。图 1 下半部分是缺陷区域凸显模块。在训练时,使用训练好的编码器作为特征提取网络,提取所有无缺陷训练样本的特征,并通过平均池化来获取最终需要保留的特征。之后,通过核心集子采样的方式构建特征库。在测试时,将测试图像通过编码器以同样的方式获取特征,并在特征库中寻找与测试图像特征最相似的特征,然后计算它们之间的欧氏距离以得到异常图。异常图在缺陷区域有较高的像素值,能粗略地凸显出缺陷区域,将其与重建差值图相乘,得到最终的融合结果。

1.2 编码器和解码器

重建网络的编码器由 4 个相同 DWTblock 构成,每个块由若干个卷积层、激活层、批量正则化层和 1 个 DWT 组成。而解码器中的 IDWTblock 由 1 个 IDWT 和多层卷积组成,具体结构如图 2 所示。其中每个卷积层都是采用卷积核大小为 3、步长为 1,扩张为 1 的卷积,即卷积不改变图像分辨率,下采样通过 DWT 完成。二维 DWT 会通过 f_{LL} , f_{LM} , f_{LH} 和 f_{HH} 这 4 个滤波器将图像分解为 1 个低频和 3 个高频,每个频率分量宽高为原来 1/2,而这 4 个滤波器分别由 LL^T , HL^T , LH^T 和 HH^T 得到,其中 $L^T = \frac{1}{\sqrt{2}}[1, 1]$, $H^T = \frac{1}{\sqrt{2}}[1, -1]$ 。

1.3 自注意力模块

自注意力机制是一种在深度学习中广泛应用的机

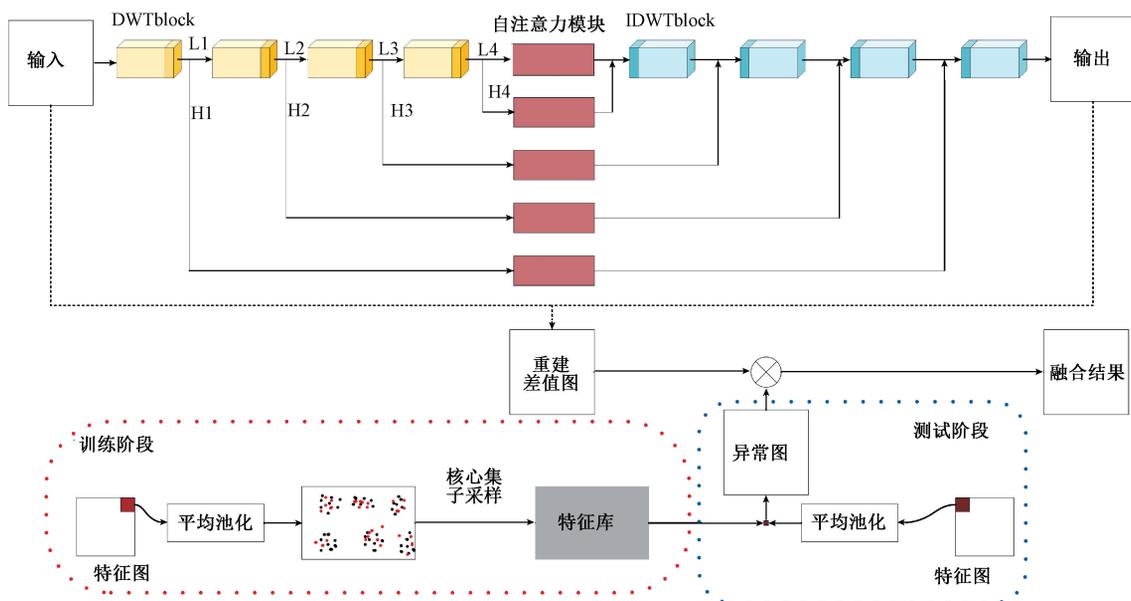


图 1 整体结构

Fig. 1 The architecture of proposed method

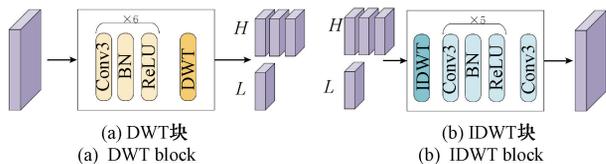


图 2 DWT 块和 IDWT 块结构图

Fig. 2 Structure of DWT block and IDWT block

制,可以学习到输入序列中不同部分之间的相互依赖关系。注意力机制的核心是通过网络把权重聚集在图像的重要区域而忽视不相关的局部。具体公式如下:

$$\text{attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{d^k}\right) \mathbf{V} \quad (1)$$

式中: \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 分别为查询向量、键向量和值向量, d^k 为向量 \mathbf{K} 的维度。通过将自注意力模块加入到跳连接中,希望编码器在向解码器传递信息时,保留正常区域的特征,同时忽视缺陷区域的特征,以此来抑制异常区域的重建。

1.4 数据增强

为了增加样本的多样性,选择用 CutPaste^[16] 的方式在训练阶段做一种数据增强。简单来说,就是对原图选择一定大小、一定长宽比的矩形区域做剪切和粘贴。CutPaste 有两种形式,一种是对长方形区域进行裁剪,另一种是对细长区域进行裁剪,类似于划痕。此外,还添加了 Cutout 的模式,即随机裁剪。具体效果如图 3 所示。其中,矩形区域的宽高、位置和角度都是随机生成的。在训练时,输入图片会随机选取一种数据增强方式,或者直

接不经过数据增强送入网络,充分保证了数据的随机性。通过这种数据增强方式,可以产生一定的空间不规则性,作为没有参与训练的真实缺陷样本的近似。虽然这种方法不能完美地模拟真实的缺陷,但可以在一定程度上与正常样本的模式区分开,使得网络不仅能够学习从正常到正常的模式,还可以学习到从非正常到正常的模式。同时,这种方式还能一定程度上解决过拟合问题,增强网络的鲁棒性。

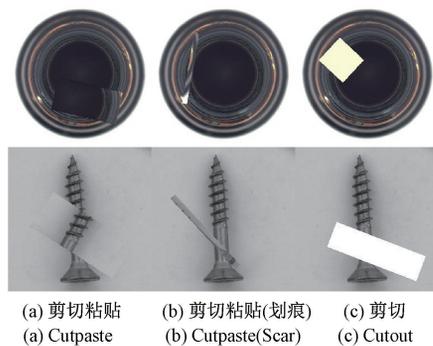


图 3 数据增强示例图

Fig. 3 Examples of data augmentation

1.5 重建网络的损失函数和异常评分

L2 损失只关注两个单独像素点之间的差异,忽略了像素之间的联系,只适用纹理均匀的图像重构,如布匹类图像。而对于像瓶子这类物体类图像,单纯的 L2 损失不能满足对复杂纹理和边缘重建的要求。因此,还引入了结构相似性损失,结构相似性(structural similarity, SSIM)

是一种用于捕捉目标感知相似性的距离度量,它捕获局部像素区域之间的相互依赖关系,更关注整体亮度、对比度和结构上面的差异。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

$$L_{ssim}(x, y) = 1 - SSIM(x, y) \quad (3)$$

$$L_{total} = L_2(x, \hat{x}) + \alpha L_{ssim}(x, \hat{x}) \quad (4)$$

其中, μ 为均值, σ 为方差, c_1 和 c_2 为常数项。 x 为输入, \hat{x} 为重建结果。 α 为 0~1 之间的损失加权系数,对于纹理图像取 0,对于物体类图像取 1。

将重建图和原图求差值并按通道取均值,再经过高斯平滑得到差值图,用来计算重建网络像素级结果。为了避免在重建过程中背景和噪声的影响,采用重建差值图中所有像素的最大值作为图像级的异常评分。

1.6 缺陷区域凸显模块

由于网络重建能力的限制和噪声的影响,会在非缺陷区域产生一定的重建误差。且部分缺陷像素与正常像素相似,导致在求差值时,这些区域无法通过阈值将其与重建不准确的正常区域区分开,从而影响像素级缺陷检测的结果。为了加强缺陷区域的差异,本文提出了一个缺陷区域凸显模块,具体结构如图 1 下半部分。假设得到的特征图大小为 $H \times W$,用 3×3 的窗口对特征图进行滑动,滑动步长为 1,特征图扩张为 1。最终会得到 $H \times W$ 个 patch,经过平均池化后就是要保留的特征。为了降低所需要存储特征的数量,受到 PatchCore^[14] 的启发,使用了核心集子采样的方法对所有正常样本的特征构建了一个特征库,特征库的大小为正常样本特征数量的 1/100 左右。采样公式如下:

$$m_i \leftarrow \underset{m \in M - M_C}{\operatorname{argmax}} \underset{m_c \in M_C}{\min} \|m - m_c\|_2 \quad (5)$$

其中, M 为所有正常样本特征的集合; M_C 为核心集子采样得到的特征库,初始为空。 m_i 表示第 i 次采样的结果,每次采样都会将 m_i 送入 M_C 中。

$$m^* = \underset{m \in M_C}{\operatorname{argmin}} \|m^{test} - m\|_2 \quad (6)$$

$$s = \|m^{test} - m^*\|_2 \quad (7)$$

对于测试集,以同样的方式提取 patch 级别的特征 m^{test} ,然后在特征库中找到与其最相似的正常样本特征 m^* ,并计算两者之间的欧氏距离,以得到每个图像块的异常分数 s 。最后得到 $H \times W$ 大小的异常分数图,将其上采样到原图大小就得到了异常图。异常图在 patch 级别的特征层面上显示了测试图像与正常图像的差异,在缺陷区域有较高的异常分数,能够帮助凸显重建差值图中差异不明显的缺陷区域。将异常图与重建差值图相乘后的结果作为优化后的缺陷定位结果,用于最终像素级缺陷检测。

考虑到网络不同深度的特征所包含的信息不同,在实际应用中,需根据图像所属的类别选择合适的特征来得到异常图。浅层特征包含纹理、位置和颜色等基础特征,深层特征包含更多的上下文语义信息。对于有固定结构的物体类数据,如杯子,由于它们包含更多的结构和上下文信息,因此使用网络的后 3 层特征比较合适。对于比较规则且主要包含纹理信息的纹理类数据,如纱布,由于它们主要包含纹理和颜色信息,所以只需使用网络的第 1 层特征。

2 实验结果与分析

2.1 数据集与实验设置

本文采用的数据集是 MVTEC AD^[17],该数据集是无监督缺陷检测中比较有挑战性的数据集,包含 10 种物体类型和 5 种纹理类型的对象。训练集中只有无标签的正常数据,测试集包含正常和异常数据以及对应的标签。其中图像分辨率在 700×700 和 1024×1024 之间,在实验中统一缩放到 256×256 的大小。实验在 Ubuntu 操作系统下进行,采用 RTX3090 训练,初步训练轮数为 2 000 轮,使用 Adam 优化器,学习率设置为 10^{-4} 。本文采用的评价指标是可以通过设置阈值来检测的受试者工作特性下的面积 (area under the receiver operating characteristic curve, AUROC)。AUROC 是目前无监督缺陷检测中最为通用的评价指标,其取值在 0~1 之间,值越大性能越好。在实验中,计算图像级和像素级两个指标。

为了验证所提方法的性能,会与其他无监督缺陷检测方法比较,并进行了消融实验验证所提方法各个模块的有效性。

2.2 实验结果

在图 4 中展示了本文方法的效果图,从图中可以发现重建网络可以较为准确地定位到缺陷出现的位置。但对于一些图像,由于部分缺陷与正常区域相似,仅靠差值不能很好的将它们与重建不准确的正常区域区分开。而异常图在缺陷区域有较高的异常值,但它们的边界比较粗糙,而且对于部分数据其结果并不准确,如第 3 行的金属螺母图像。综合考虑,将重建结果与异常图结合,使缺陷区域被进一步凸显,同时缺陷定位更为准确。

同时在本节中,将会与不同的无监督缺陷检测的方法相比较,包括基于重建的方法 DRAEM^[18]、OCRGAN^[10]、InTra^[19],基于特征嵌入的方法 Padim^[13]、CSFLOW^[20]、PatchCore^[14],具体结果如表 1 所示。根据表 1 中的数据,本文方法的图像级和像素级缺陷检测结果均优于所提到的无监督缺陷检测算法,显示了算法的有效性。

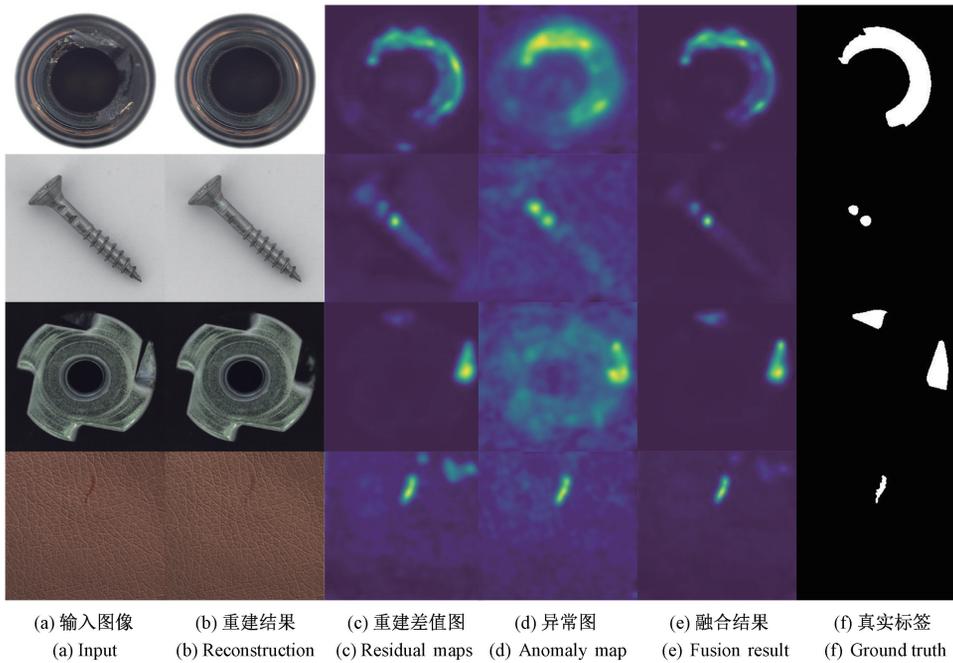


图4 重建网络和缺陷区域凸显模块结果

Fig. 4 Results of reconstruction network and defect highlighting module

表1 在 MVTec AD 数据集上图像级(左)和像素级(右) AUROC 结果

Table 1 Image level (left) and pixel level (right) AUROC results on the MVTec AD dataset

(%)

类别	基于重建的方法			基于特征的方法			本文方法 仅重建	本文方法
	DRAEM	OCRGAN	InTra	Padim	CSFLOW	PatchCore		
Carpet	97.0/95.5	99.4/—	98.8/ 99.2	—/99.1	100 /—	98.7/99.0	97.4/96.9	97.4/97.2
Grid	99.9/ 99.7	99.6/—	100 /98.8	—/97.3	99.0/—	98.2/98.7	100 /98.1	100 /98.6
Leather	100 /98.6	97.1/—	100 / 99.5	—/99.2	100 /—	100/99.3	100 /98.9	100 /99.4
Tile	99.6/ 99.2	95.5/—	98.2/94.4	—/94.1	100 /—	98.7/95.6	100 /98.1	100 /98.0
Wood	99.1/96.4	95.7/—	97.5/88.7	—/94.9	100 /—	99.2/95	100 /96.9	100 / 97.8
Bottle	99.2/ 99.1	99.6/—	100 /97.1	—/98.3	99.8/—	100 /98.6	100 /98.6	100 /99.0
Cable	91.8/94.7	99.1/—	70.3/91.0	—/96.7	99.1/—	99.9 / 98.4	98.6/97.5	98.6/98.3
Capsule	98.5 /94.3	96.2/—	86.5/97.7	—/98.5	97.1/—	98.1/ 98.8	97.3/98.4	97.3/ 98.8
Hazelnut	100 / 99.7	98.5/—	95.7/98.3	—/98.2	99.6/—	100 /98.7	100 /98.1	100 /98.6
Metal nut	98.7/ 99.5	99.5/—	96.9/93.3	—/97.2	99.1/—	100 /98.4	99.5/97.7	99.5/98.2
Pill	98.9/97.6	98.3/—	90.2/98.3	—/95.7	98.6/—	96.6/97.4	99.6 /98.1	99.6 / 98.4
Screw	93.9/97.6	100 /—	95.7/ 99.5	—/98.5	97.6/—	98.1/99.4	98.4/94.7	98.4/97.8
Toothbrush	100 /98.1	98.7/—	100/98.9	—/98.8	91.9/—	100 /98.7	100 / 99.4	100 /99.3
Transistor	93.1/90.9	98.3/—	95.8/96.1	—/ 97.5	99.3/—	100 /96.3	99.7/96.6	99.7/97.0
Zipper	100 /98.8	99.0/—	99.4/99.2	—/98.5	99.7/—	99.4/98.8	100/ 99.4	100 /99.3
Average	98.0/97.3	98.3/—	95.0/96.6	97.9/97.5	98.7/—	99.1/98.1	99.3 /97.8	99.3 / 98.3

2.3 消融实验

为了证明所提出方法各个模块的有效性,分别对重建网络和缺陷区域凸显模块进行消融实验。

为了验证自注意力模块的有效性,将重建网络中的自注意力模块替换为普通的记忆模块进行实验,并在

图5展示了两种方法的重建效果,同时表2中记录了两种重建网络缺陷检测图像级和像素级的结果。

由于自注意力更加关注全局的感知,能更好地重建出与周围不一致的缺陷区域,所以重建效果要更好。同时表2中的结果也表明,本文方法的重建网络相比于添

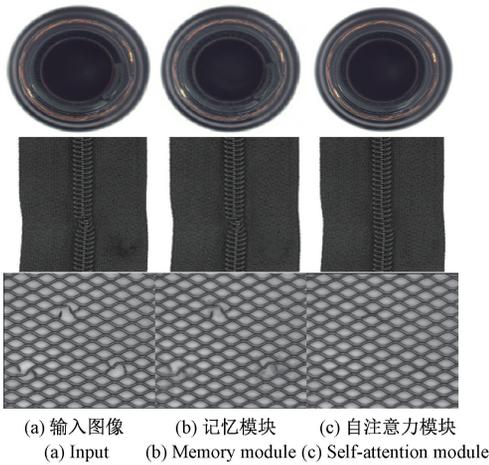


图 5 记忆模块和自注意力模块的重建结果

Fig. 5 Reconstruction results with memory module and self-attention module

加记忆模块的网络具有更好的缺陷检测和定位性能,从而验证了自注意力模块的优势。

表 2 记忆模块和自注意力模块的重建网络对比

Table 2 Comparison of reconstruction networks with memory module and self-attention module (%)

类别	Memory 模块	自注意力模块
Carpet	99.1/96.8	97.4/96.9
Grid	97.0/95.2	100/98.1
Leather	98.5/98.4	100/98.9
Tile	92.3/96.3	100/98.1
Wood	100/95.4	100/96.9
Bottle	99.1/95.5	100/98.6
Cable	98.0/ 97.7	98.6/97.5
Capsule	86.0/97.5	97.3/98.4
Hazelnut	97.8/ 98.8	100/98.1
Metal nut	97.8/96.2	99.5/97.7
Pill	94.5/95.2	99.6/98.1
Screw	90.5/84.8	98.4/94.7
Toothbrush	100/98.6	100/99.4
Transistor	98.7/85.4	99.7/96.6
Zipper	97.7/98.2	100/99.4
Average	96.4/95.3	99.3/97.8

为了验证在重建网络中采用 DWT 的效果,将网络中 DWT 替换成池化,并将 IDWT 替换成上采样,保留网络其他部分不变。具体结果如表 3 所示。相较于传统的重建网络,基于 DWT 的网络在图像级 AUROC 和像素级 AUROC 的结果分别提升了 1.3% 和 1.2%。

同时,对于图像分辨率的选择也进行了实验,实验随机选取 bottle、cable、grid 共 3 类,记录了重建网络的结果,具体数据如表 4 所示。由于重建网络包含多个自注

意力模块,图像分辨率每提高一点,所需的显存大大增加,综合性能和计算成本考虑,选择 256 的分辨率最为合适。

表 3 基于 DWT 和池化的重建网络对比

Table 3 Comparison of reconstruction networks with DWT and pooling (%)

类别	池化+上采样	DWT
Carpet	93.1/92.6	97.4/96.9
Grid	98.5/97.3	100/98.1
Leather	97.4/97.0	100/98.9
Tile	99.6/96.8	100/98.1
Wood	100/97.5	100/96.9
Bottle	99.8/97.5	100/98.6
Cable	96.8/96.1	98.6/97.5
Capsule	95.6/97.6	97.3/98.4
Hazelnut	99.2/ 99.3	100/98.1
Metal nut	98.4/97.3	99.5/97.7
Pill	95.7/97.1	99.6/98.1
Screw	96.4/ 94.7	98.4/94.7
Toothbrush	100/99.3	100/99.4
Transistor	99.7/91.2	99.7/96.6
Zipper	99.8/99.1	100/99.4
Average	98.0/96.6	99.3/97.8

表 4 不同分辨率的影响

Table 4 The impact of different resolutions

分辨率	128	224	256	288
显存/G	3.6	10.9	16.5	26.7
结果/%	98.0/95.9	98.7/98.1	99.5/98.0	99.6/98.0

对于缺陷区域凸显模块,从表 1 中的结果可以发现,相较于单独的重建网络,重建网络结合缺陷区域凸显模块的方法得到的像素级缺陷检测结果有明显的提升。为了验证不同层的特征经过缺陷区域凸显模块的效果,进行了单层特征和多层特征拼接融合的实验。为了保证实验的公平性,对特征库的子采样倍率统一设为 0.01。考虑到第 1 层特征维度较大,生成的特征库占用空间较大,因此没有考虑将第 1 层和后面的特征结合。表 5 和 6 分别展示了物体类和纹理类图像的像素级 AUROC 结果。第 1 行显示的是单独使用缺陷区域凸显模块得到的结果,而第 2 行则表示本文所用方法,即重建差值图和异常图相结合后的结果。

表 5 物体类数据不同特征层的结果

Table 5 Results of different feature layers about object data (%)

Layer	1	2	3	4	2+3+4
缺陷区域凸显模块	86.50	91.83	85.80	74.84	95.8
本文方法	97.76	98.17	97.92	97.23	98.27

表6 纹理类数据不同特征层的结果

Table 6 Results of different feature

layers about texture data (%)

Layer	1	2	3	4
缺陷区域凸显模块	95.08	83.7	58.0	53.14
本文方法	98.21	97.34	85.06	81.78

从表5中数据可以看出,对于物体类的数据,将最后3层特征进行拼接融合,并将其与重建网络的差值图结合,相较于单层特征,这种方式获得了更好的效果。而对于纹理类数据,如表6所示,只使用了第1层的特征就取得了较好的结果。为了减小使用第1层特征生成的特征库的维度,需要将特征库的子采样倍率设定得更低。针对特征核心集子采样倍率的选择问题,也进行了一系列实验,对应像素级 AUROC 结果如表7所示。从表7中数据综合考虑,对于物体类数据,特征库的大小为原来的所有特征的0.05比较合适,而纹理类数据选择0.005更合适。

表7 不同采样倍率的影响

Table 7 The effect of different

downsampling magnifications (%)

倍率	0.001	0.005	0.010	0.050	0.100
物体类		98.14	98.27	98.47	98.50
纹理类	98.10	98.20	98.20	98.20	

3 结 论

本文提出了一种将自注意力与小波相结合的重建网络。该网络能够分离高频的细节信息和低频的上下文语意信息,并使网络在传递信息时重点关注正常区域而忽视缺陷区域,提高了网络的重建能力。同时,还提出了一个缺陷区域凸显模块,用于提供额外的缺陷位置信息,以改善重建网络重建不准确的问题,从而提高缺陷定位能力。该方法在MVTec AD数据集上取得了较好的效果,优于所提到的对照方法。然而,这两项改进也带来了一定的问题,如网络规模相较于以前的重建网络变大,因此未来的改进方向是如何使其变得更加轻量级。

参考文献

- [1] CHEN Y, DING Y, ZHAO F, et al. Surface defect detection methods for industrial products: A review[J]. Applied Sciences, 2021, 11(16): 7657.
- [2] 赵朗月, 吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2022, 43(1): 198-219.
- ZHAO L Y, WU Y Q. Research progress of surface defect detection method based on machine vision [J].

Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.

- [3] TAO X, GONG X, ZHANG X, et al. Deep learning for unsupervised anomaly localization in industrial images: A survey[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-21.
- [4] 辛彦, 吴国新, 左云波. 基于EfficientDet的风机叶片缺陷检测方法[J]. 电子测量技术, 2022, 45(5): 124-131.
- XIN Y, WU G X, ZUO Y B. Defect detection method of wind turbine blade based on EfficientDet[J]. Electronic Measurement Technology, 2022, 45(5): 124-131.
- [5] 单明陶, 高玮玮, 杨亦乐, 等. 基于改进Faster R-CNN的刹车零件缺陷检测算法[J]. 国外电子测量技术, 2022, 41(4): 22-28.
- SHAN M T, GAO W W, YANG Y L, et al. Defect detection algorithm of brake parts based on improved Faster R-CNN [J]. Foreign Electronic Measurement Technology, 2022, 41(4): 22-28.
- [6] ZHOU C, PAFFENROTH R C. Anomaly detection with robust deep autoencoders[C]. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017: 665-674.
- [7] AKCAYK S, ATAPOUR-ABARGHOU EI A, BRECKON T P. Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection [C]. 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-8.
- [8] GONG D, LIU L, LE V, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1705-1714.
- [9] PARK H, NOH J, HAM B. Learning memory-guided normality for anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 14372-14381.
- [10] LIANG Y, ZHANG J, ZHAO S, et al. Omni-frequency channel-selection representations for unsupervised anomaly detection [J]. arXiv preprint arXiv: 2203.00259, 2022.
- [11] WU K, ZHU L, SHI W, et al. Self-attention memory-augmented wavelet-CNN for anomaly detection [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(3): 1374-1385.
- [12] COHEN N, HOSHEN Y. Sub-image anomaly detection with deep pyramid correspondences [J]. arXiv preprint arXiv: 2005.02357, 2020.

- [13] DEFARD T, SETKOV A, LOESCH A, et al. PaDiM: A patch distribution modeling framework for anomaly detection and localization [C]. International Conference on Pattern Recognition. Springer, Cham, 2021, 12664: 475-489.
- [14] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14318-14328.
- [15] 李亚舟, 曹江涛, 姬晓飞. 融合改进 Padim 建模和 ResNet 网络的喷涂质量检测算法 [J]. 电子测量与仪器学报, 2022, 36(11): 91-97.
LI Y ZH, CAO J T, JI X F. Spraying quality detection algorithm by fusing improved Padim modeling with ResNet network [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(11): 91-97.
- [16] LI C L, SOHN K, YOON J, et al. Cutpaste: Self-supervised learning for anomaly detection and localization [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 9664-9674.
- [17] BERGMANN P, FAUSER M, SATTLEGGER D, et al. MVTec AD-A comprehensive real-world dataset for unsupervised anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9592-9600.
- [18] ZAVRTANIK V, KRISTAN M, SKOCAJ D. Draem-A discriminatively trained reconstruction embedding for surface anomaly detection [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 8330-8339.

- [19] PIRNAY J, CHAI K. Inpainting transformer for anomaly detection [C]. International Conference on Image Analysis and Processing. Lecture Notes in Computer Science. Springer, Cham, 2022, 13232: 394-406.
- [20] RUDOLPH M, WEHRBEIN T, ROSENHAHN B, et al. Fully convolutional cross-scale-flows for image-based defect detection [C]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 1088-1097.

作者简介



乐华钢, 2020 年于武汉科技大学获得学士学位, 现为武汉科技大学硕士研究生, 主要研究方向为计算机视觉和缺陷检测。
E-mail: 1060935223@qq.com

Yue Huagang received his B. Sc. degree from Wuhan University of Science and Technology in 2020. Now he is a M. Sc. candidate in Wuhan University of Science and Technology. His main research interests include computer vision and defect detection.



王文武 (通信作者), 2013 年于华中科技大学获得博士学位, 现为武汉科技大学副教授, 主要研究方向为计算机视觉和模式识别。
E-mail: wangwenwu@wust.edu.cn

Wang Wenwu (Corresponding author) received his Ph. D. degree from the Huazhong University of Science and Technology in 2013. Now he is an associate professor in Wuhan University of Science and Technology. His main research interests include computer vision and pattern recognition.