

基于改进自编码器结构的轮胎缺陷检测*

李洪奎 陈 浩 刘韵婷 张兴伟 冯欣悦

(沈阳理工大学自动化与电气工程学院 沈阳 110159)

摘要:针对部分轮胎X光缺陷图像中缺陷对比度较低、缺陷占比较小,导致缺陷难以检测的问题,采用了一种基于生成对抗网络的改进模型,以提高轮胎缺陷的检测精度。首先分析了传统生成器所存在的一些问题,然后以GANomaly作为基础模型,引入了注意力机制模块NAM、流对齐模块FAM和PatchGAN,旨在增强模型的特征提取能力和图像重构能力。注意力机制模块NAM通过归一化处理增强了模型对缺陷区域的关注度,流对齐模块能够将低分辨率特征图中的特征点精确地映射到高分辨率特征图的对应位置,从而确保多尺度特征之间的信息一致性和有效融合,而PatchGAN则通过局部判别器增强了模型对局部特征的识别能力。为了验证改进模型的有效性,在相同的自制数据集上对4种轮胎缺陷类型X光图片进行测试。测试结果表明,改进后的模型在受试者工作特征曲线面积(AUC)和平均精度(AP)两个关键指标上均取得了显著提升,AUC值达到了96.4%,AP值达到了95.8%。这些结果表明,改进后的模型有效增强了特征提取和图像重构的能力,提升了缺陷检测的精准度。

关键词:生成对抗网络;NAM;深度学习;FAM;轮胎缺陷检测;PatchGAN

中图分类号:TP274;TN911

文献标识码:A

国家标准学科分类代码:510.4

Tire defect detection based on improved autoencoder structure

Li Hongkui Chen Hao Liu Yunting Zhang Xingwei Feng Xinyue

(School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract: To address the challenges of low contrast and small defect sizes in some X-ray images of tires, which make detection difficult, an improved model based on generative adversarial networks (GANs) is proposed to enhance the accuracy of tire defect detection. Initially, issues with traditional generators are analyzed. Building upon the GANomaly model, the proposed approach incorporates the attention mechanism module (NAM), flow alignment module (FAM), and PatchGAN to enhance feature extraction and image reconstruction capabilities. The NAM enhances the model's focus on defect areas through normalization, while the FAM accurately maps features from low-resolution to high-resolution feature maps, ensuring information consistency and effective fusion across multiple scales. PatchGAN, with its local discriminator, improves the model's ability to recognize local features. Validation tests on a self-constructed dataset of four tire defect types demonstrate significant improvements in key metrics, achieving an AUC of 96.4% and an AP of 95.8%. These results indicate enhanced feature extraction and image reconstruction capabilities, leading to improved defect detection accuracy.

Keywords: generative adversarial network; NAM; deep learning; FAM; tire defect detection; PatchGAN

0 引言

在当今的汽车工业中,子午线轮胎已经成为了主流的轮胎类型,因其优异的操控性能和乘坐舒适性,深受广大消费者的喜爱。然而,子午线轮胎在生产过程中可能

产生多种缺陷,如气泡、裂纹、杂质等,这些缺陷如果不被及时发现和处理,可能会引发严重的安全问题。

随着深度学习的发展,无监督算法^[1-4]在异常检测领域中展现出了显著优势,基于无监督算法的二分类模型对于新的、未见过的异常类型有更好的泛化能力。在实际应用中,可能会遇到之前未见过的缺陷类型,其仅需使

用正常类型的图片进行训练即可进行缺陷检测。无需像有监督算法^[5-8]那种多分类模型依赖大量带有标注且预先定义缺陷类型的缺陷轮胎图像进行训练。基于有监督算法这种多分类模型,如果标注错误,模型可能会学习到错误的特征,导致误判。相比之下,二分类模型则不存在这些问题。这种优势使得在轮胎生产等领域,选择二分类模型能够可靠地确保产品质量和生产效率,避免由于新类型异常的出现和缺陷标注错误所带来的不良影响。

生成对抗网络 (generative adversarial network, GAN)^[9] 作为一项关键的无监督学习技术,已被广泛应用于异常检测领域。其核心思想是利用生成的重构数据与原始输入数据之间的差异性来判断被检测物是否有缺陷。AnoGAN^[10] 开辟了使用 GAN 进行异常检测的先河。AnoGAN 算法有两个阶段,首先利用正常数据训练一个标准的 GAN 网络,并在第 2 步中通过迭代方法将测试样本映射至隐空间以获取相应的隐码。利用这一隐码,经训练的 GAN 进而能够重构出对应的样本,将输入样本与重构样本之间的差异度量视作异常评分的标准。在此之后,各种基于 GAN 的异常检测算法陆续问世,通过加入各式各样的子网络以增强性能和精准度,例如 GANomaly^[11]、f-AnoGAN^[12] 和 OCGAN^[13]。此外还有更多研究尝试结合来自不同学术领域的理论和技术,目的是提升缺陷检测算法的能力,以解决实际应用中可能出现的各种问题。

在 2022 年,He 等^[14]提出了一种基于 GAN 的无监督缺陷检测算法,其创新之处在于实现了边缘修复和特征融合。为了进一步提升边缘修复的效果,设计了一种可变形的自动编码器。这种编码器设计利用了可变形卷积的特性,通过偏移量来调整卷积核的形状,从而更好地捕捉图像中的不规则形状和复杂结构,提升了模型的重构能力。

Li 等^[15]于 2023 年提出了一种轮胎缺陷图像生成模型。该模型将残差网络和注意力机制引入到生成器中,以提高模型的特征提取能力。并且采用带有梯度惩罚项的 Wasserstein 距离,取代了传统生成对抗网络中使用的 JS 散度损失函数,提高了模型训练的稳定性。

当前,GAN 中的生成器主要采用了基于自编码器 (auto encoder, AE)^[16] 的模型架构。AE 在进行特征提取的过程中,下采样操作会导致一些重要信息丢失,而在上采样过程中会引入位置误差,导致不同层次的特征图在空间维度上无法精确对齐,进而导致重构图像不够真实。而 GAN 原有的判别器只关注图像的整体效果,但在判别局部细节质量方面表现不足。这会导致一些局部缺陷无法被发现,影响缺陷检测效果。

为了改善以上问题,以 GANomaly 作为基础模型,在两个编码器部分集成了基于标准化的注意力模块

(normalization based attention module, NAM)^[17],旨在通过分配加权关注提升模型对输入数据中重要特征的识别力。同时,在解码器部分引入了流对齐模块 (flow alignment module, FAM)^[18],以优化重建图像的质量,确保模型能够更准确地复原出正常的数据分布。此外,本文采用 PatchGAN^[19] 判别器,借助其在局部区域的判断能力,增强了模型对细粒度异常特征的识别。通过对自制轮胎 X 光图像数据集中的 4 种缺陷进行测试,实验结果表明,相比基础模型,改进后的模型能够有效提升缺陷检测能力。

1 研究内容

1.1 NAM 注意力机制

编码器在特征提取的过程中采用下采样操作,可使特征图的尺寸减少,从而提取更高级别、更抽象的特征,同时减少计算量和参数数量。但是在这个过程中会出现信息丢失的问题。这会导致模型特征提取过程中缺少了某些重要的信息,从而影响模型的性能。引入注意力机制模块可以减少上述问题对模型的影响。

近年来,注意力机制在深度学习领域中引起了广泛的关注,因为它能够帮助神经网络专注于最关键的信息,同时抑制次要的细节。尽管许多早期研究展示了通过注意力机制来捕捉突出特征的有效性,但是忽略了权重分配在抑制不显著通道方面的潜在作用。为了优化通道注意力子模块的性能,NAM 巧妙地将批归一化的缩放因子融入权重计算过程。通过这种设计,可以为那些具有更重要的通道分配更多权重,从而使得网络在提炼和强化关键特征上更为专注。此外,相较于压缩激励 (squeeze and excitation, SE)^[20]、瓶颈注意力模块 (bottleneck attention module, BAM)^[21] 和卷积块注意力模块 (convolutional block attention module, CBAM)^[22] 等传统注意力机制,NAM 的一大优势在于它摒弃了占用大量计算资源的全连接层和卷积层,使模型更轻量化。

NAM 采用了 CBAM 的模块集成方式,该模块由注意力通道和空间注意力机制组成。针对这两个子模块,提出了创新的设计理念,使之能够灵巧地部署在网络的卷积块之后。在通道注意力子模块的设计框架内,整合了 Batch Normalization^[23] 所定义的缩放因子。该因子能够体现通道变化的程度,其具体表达如式 (1) 所示,强调那些最为重要的通道。

$$B_{out} = BN(B_{in}) = \gamma \frac{B_{in} - \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \varepsilon}} + \beta \quad (1)$$

式中: μ_{β} 作为小批量 B 的均值, σ_{β} 作为其标准差, γ 和 β 则作为可训练的仿射变换参数,分别用于尺度和偏移的

调整。

如图 1 和相应式(2)展示的通道注意力子模块定义了输出特征 M_c 利用比例因子 γ 对各通道进行调节,因此,可以得到每个通道的权值如式(3)所示。如果对空间中的每个像素使用相同的归一化方法,就可以得到空间注意力的权重,这个过程就被称为像素归一化。图 2 和式(4)详细介绍了空间注意力模块及其输出 M_s 。 λ 是缩放因子,权重如式(5)所示。NAM 的损失函数由特征预测损失和权重稀疏惩罚组成,其中 l 是特征预测的损失函数, $f(x, W)$ 表示模型的预测结果, y 是真实标签。

其中, p 是正则化系数, $g(\gamma)$ 和 $g(\lambda)$ 是关于缩放因子的惩罚项。具体函数如式(6)所示。

综上所述,在编码器部分引入 NAM 注意力机制,不

仅增强模型的特征提取能力,高效地提炼出有用的信息。还减少在不相关特征上的过度学习,降低过拟合风险。

$$M_c = \text{sigmoid}(W_\gamma(BN(F_1))) \quad (2)$$

$$\omega_i = \frac{\gamma_i}{\sum_{j=0} \gamma_j} \quad (3)$$

$$M_s = \text{sigmoid}(W_\lambda(BN_s(F_2))) \quad (4)$$

$$\omega_i = \frac{\lambda_i}{\sum_{j=0} \lambda_j} \quad (5)$$

$$\text{Loss} = \sum_{(x,y)} l(f(x, W), y) + p \sum g(\gamma) + p \sum g(\lambda) \quad (6)$$

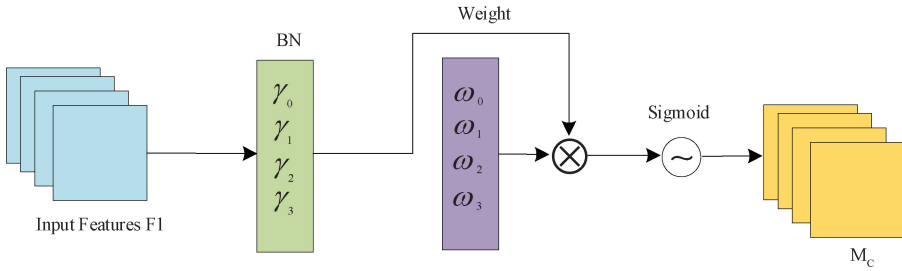


图 1 通道注意力模块

Fig. 1 Channel attention mechanism

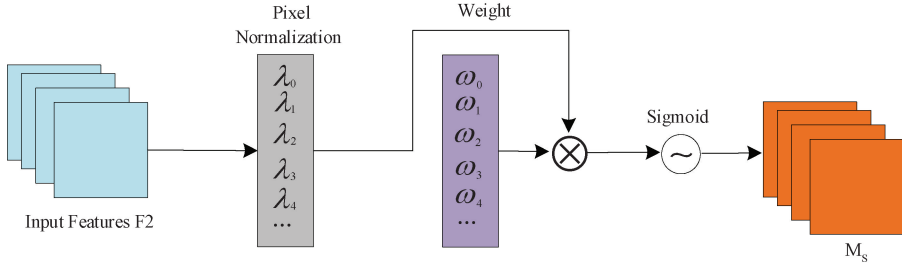


图 2 空间注意力模块

Fig. 2 Spatial attention mechanism

1.2 流对齐模块

在特征提取过程中,下采样操作会导致特征图中的位置信息丢失,多个相邻像素会被映射到同一个较大的像素位置上。在上采样时,由于无法准确还原这些丢失的位置信息,可能会导致特征图的空间上不匹配的问题。

因此,选择在解码器部分引入流对齐模块,将解码器中低分辨率特征图 and 高分辨率特征图经过流对齐模块生成一个个新的特征图,然后将这些特征图进行融合。通过这种方式,可以把低分辨率特征图上的每一个特征点精确地映射到高分辨率特征图上的对应位置,从而解决了不同层级特征图之间的对齐问题。这种方法在充分利

用不同层次之间特征的同时,保证信息的准确对齐,提升模型的重构性能。

给定两个不同层级不同分辨率的特征图 F_L 和 F_{L-1} 。利用流对齐模块,将 F_L 的分辨率通过上采样调整至与 F_{L-1} 相同。接下来,将调整后的 F_L 特征图与 F_{L-1} 特征图合并,并进一步通过一个含有两个尺寸为 3×3 的卷积核的卷积层进行处理。就能生成一个语义流场(semantic flow field),流对齐模块的输出是语义流场的预测值。其数学过程如式(7)所示。

$$\Delta_{L-1} = \text{conv}_L(\text{cat}(F_L, F_{L-1})) \quad (7)$$

其中, $\text{cat}(\cdot)$ 表示 Concat 拼接操作,而 $\text{conv}_L(\cdot)$ 是 3×3 卷积层。流对齐模块如图 3 所示。

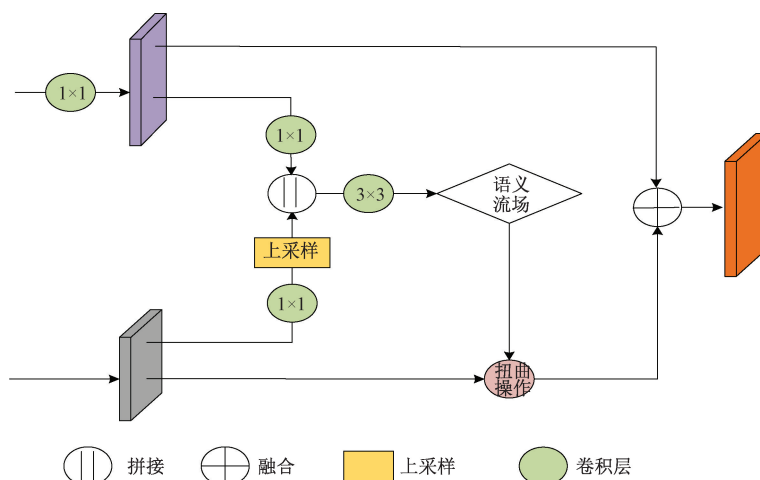


图 3 流对齐模块

Fig. 3 Flow alignment module

在语义流场中空间网格 Ω_{L-1} 内的任意一个点 P_{L-1} ，都能够通过简单的加法运算映射至与其对应的上采样位置 P_L 。通过观察图 4，可以看出特征图与流场呈现出分辨率上的不同。所以应用式 (8) 所描述的方法，对其进行减半处理，纠正这种差异。然后，借助空间变换网络 (spatial transformer networks, STN)^[24] 中提及的微分双线性采样策略，针对 P_L 的 4 个周边点 (即左上、右上、左下、

右下) 进行采样，并通过线性插值合并，目的是精确地计算出最终的语义流场映射。其数学过程如式 (9) 所示。

$$P_L = \frac{P_{L-1} + \Delta_{L-1}(P_{L-1})}{2} \quad (8)$$

$$\tilde{F}_L(P_{L-1}) = F_L(P_L) = \sum_{P \in N(P_L)} \omega_P F_L(P) \quad (9)$$

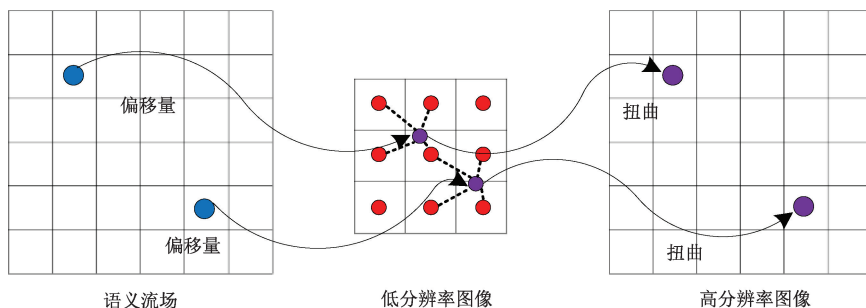


图 4 扭曲过程

Fig. 4 Warp Procedure

其中, $N(P_L)$ 表示 F_L 中扭曲点 P_L 的邻域, ω_P 表示通过翘曲网格的距离估计的双线性核权重。这个过程在一定程度上与可变形卷积网络 (deformable convolutional networks, DCN)^[25] 的可变形内核操作相似。尽管有相似之处,但是,此方法多个方面与 DCN 不同。首先, FAM 通过预测偏移场来整合高层和低层特征,使得两者在位置上能够对齐。而 DCN 通过根据预测的位置偏移量调整内核的位置,以自适应地调整其感受野的大小。其次,流对齐模块旨在精确地对齐不同层次的特征,而 DCN 的核心功能则更偏向识别并专注于目标对象的显著部分,这可被视为一种注意力机制。

1.3 PatchGAN 网络

判别器的功能是识别由生成器重构的图像与真实图像之间的差异,并判断图像是否为伪造。传统的全局判别器网络会输出一个单一的标量值,这个值反映了输入图像被视为真实图像的可能性,并以此来评价整个图像的真实性。但这种方法存在不足之处:它输出的单一标量值是对整张图片的全局评估,无法细致反映各个局部区域的特征。

为了改善这一问题,引入 PatchGAN 网络代替原有的全局判别器。它能够更加精细地捕捉图像中的局部纹理细节,并提升对生成图片的判别能力。PatchGAN 模型

将输入映射为一个 $N \times N$ 的矩阵,矩阵中的每个元素都代表一个特定的图像块。通过计算每个感受野的平均值,可以得出最终图像为真实图像的可能性。这样可以使判

别器具有更细粒度的判别能力,可以更好地捕捉到生成图像与真实图像之间的局部差异,从而提高缺陷检测的准确性。PatchGAN 网络结构图如图 5 所示。

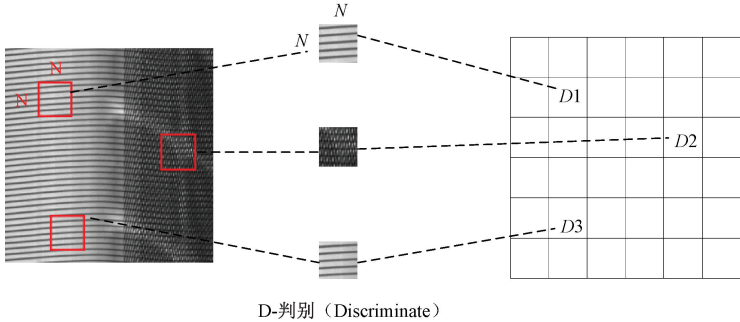


图 5 PatchGAN 网络结构

Fig. 5 PatchGAN network structure

1.4 整体网络框架

AFPGAN 的模型架构如图 6 所示,该网络模型由生成器和判别器构成。生成器部分由编码器-解码器-编码器组成,其中两个编码器整体结构是相同的,编码器部分中的第一层由卷积层 (Conv2d) 和 Leakyrelu 激活函数组成,2~6 层都依次由 Conv2d、批归一化层 (Batchnorm)、Leakyrelu 激活函数和 NAM 注意力机制组成。最后经过一层 Conv2d 转换为潜在空间向量 (Latent)。解码器部分

中第一层为转置卷积层 (ConvTranspose2d),2~6 层都依次由 ConvTranspose2d、Batchnorm 和 Relu 激活函数组成,而最后一层由 ConvTranspose2d 和 Tanh 激活函数组成。判别器部分则由 6 个层级构成,第一层由 Conv2d 和 Leakyrelu 激活函数组成,2~5 层依次由 Conv2d、Batchnorm、Leakyrelu 激活函数组成,最后一层为 Conv2d 用于分类。

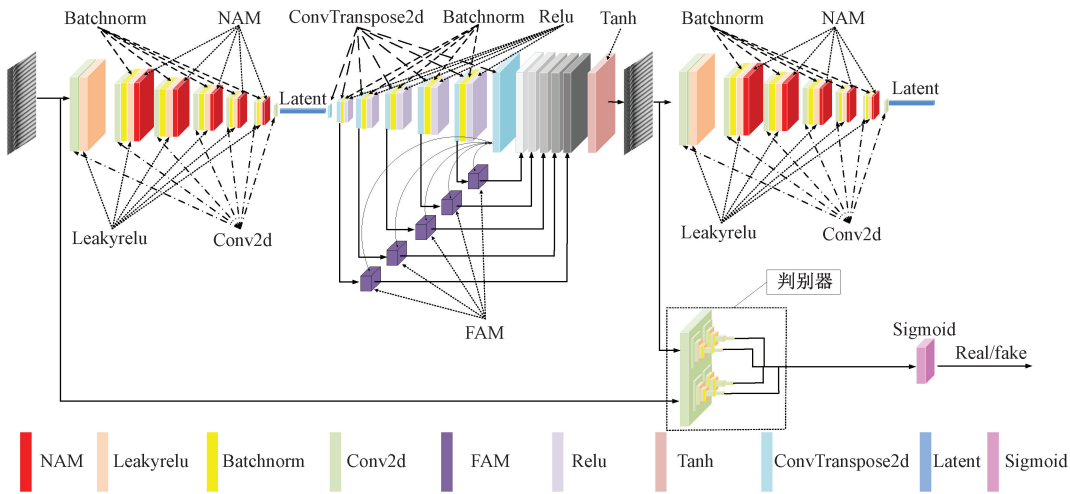


图 6 AFPGAN 网络模型

Fig. 6 AFPGAN network model

AFPGAN 网络模型采用了将 NAM 和 FAM 分别植入编码器和解码器的创新策略。在编码器中将 NAM 整合至每个卷积层的后方提高了模型在特征提取能力。而在解码器部分,运用 FAM 处理低分辨率与高分辨率特征图,将不同层级的特征图经过流对齐模块生成一个个新的特征图,再将这些特征图进行融合。有效改善了传统上采样过程中不同层级特征对不齐的问

题。此外,采用 PatchGAN 作为判别器,优化了模型的图像生成质量,这也间接提升了模型在缺陷检测任务中的表现。

1.5 损失函数

1) 生成器损失函数

生成器的损失函数包括 3 个,分别为对抗损失 L_{adv} 、

重构造误差损失 L_{con} 和编码损失 L_{enc} 。具体公式如下：

(1) 对抗损失

使用的是特征对齐的损失函数,根据输入 x 的值,选择判别器中的一个中间层级,进而计算相应的生成器层次产生的 L2 距离。对抗性损失的计算过程已在公式(10)中给出说明。

$$L_{adv} = E_{x \sim p_x} \| D(x) - D(G(x)) \|_2 \tag{10}$$

其中, $D(x)$ 判别器在真实数据样本 x 上的输出, $D(G(x))$ 为判别器在生成数据样本 $G(x)$ 上的输出, $\| \cdot \|_2$ 表示 L2 范数。

(2) 重构损失

通过计算输入图片和重构图片之间的 L1 距离,使得重构后的图片尽可能接近输入图片。重构损失函数如式(11)所示。

$$L_{con} = E_{x \sim p_x} \| x - G(x) \|_1 \tag{11}$$

公式中 x 表示原始输入图像的数据, $G(x)$ 表示重建图像的数据, $\| \cdot \|_1$ 表示 L1 范数。

(3) 编码损失

上面的两个损失函数可以让重构后的图像尽量真实,引入编码损失以最小化输入图片的编码特征 Z 与重构图像的编码特征 Z^1 之间的距离。编码损失函数如式(12)所示。

$$L_{enc} = E_{x \sim p_x} \| Z - Z^1 \|_2 \tag{12}$$

(4) 生成器总损失函数

如公式(13)所示。

$$L = \omega_{con} L_{con} + \omega_{adv} L_{adv} + \omega_{enc} L_{enc} \tag{13}$$

其中, ω_{con} , ω_{adv} , ω_{enc} 是 3 个损失函数对应的权重参数。

2) 判别器损失函数

PatchGAN 将输入图像分成 N 个不重叠的子块,然后独立地对每个子块进行真伪辨别。相应的损失函数已在式(14)中给出。

$$L_D = E_{x \sim p_x} [\log D(x_i)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \tag{14}$$

其中, D 为判别器、 E 为编码器、 G 为生成器、 Z 为输入图片的潜在空间向量、 x 为真实数据样本。

2 实验

2.1 实验数据集

1) 训练数据集

为了模型可以学习到足够的特征。本文选用 8 340 张不同位置的轮胎 X 光无缺陷图像作为训练数据集。如图 7 所示。

2) 测试数据集

测试数据集由无缺陷图像和缺陷图像组成。其中无

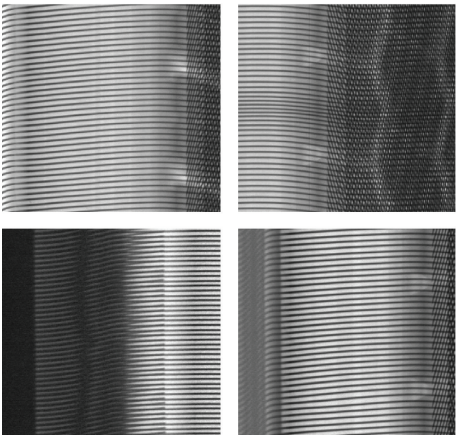


图 7 轮胎 X 光无缺陷图像

Fig. 7 Tire X-ray normal image

缺陷图像 1 900 张,缺陷图像 1 900 张。具体类型和数量如表 1 和图 8 所示。

表 1 缺陷类型及数量

Table 1 Defect type and quantity

缺陷类型	图像数量
杂质	510
气泡	500
帘线弯曲	440
侧壁裂缝	450

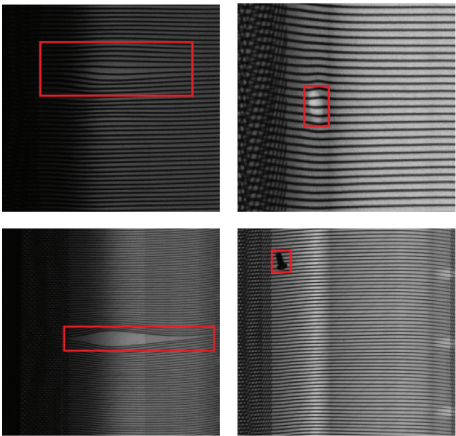


图 8 轮胎 X 光缺陷图像

Fig. 8 Tire X-ray abnormal image

2.2 实验平台配置

本次实验的开发平台配置为 Python3.9 和 Pytorch1.12.1,深度卷积神经网络开发工具. 计算机的硬件配置为 Intel(R) Xeon(R) Platinum 8260C、NVIDIA GeForce RTX 3090 GPU、86 G 内存。编程环境为 Ubuntu20.04 操作系统。

2.3 实验相关参数

AFPGAN 模型的具体参数如表 2 所示。

表 2 参数设置

Table 2 Parameter settings

参数类型	参数设置
Image size	256×256
优化器	Adam
Learning rate	0.000 3
b1	0.5
b2	0.999
epoch	100
潜在空间向量	512

2.4 实验结果分析

在本次实验中,采用了两种关键的评估工具,受试者工作特征曲线(receiver operating characteristic curve, ROC)和精确率-召回率曲线(precision recall curve, PRC)来综合评估异常检测模型的性能。这两种方法在二分类问题中具有广泛的应用,并且各有其独特的优势。

ROC 曲线通过绘制真正类率与假正类率之间的关系,直观地展示了分类器在不同决策阈值下的性能变化。而 PRC 则专注于精确率和召回率的平衡,特别是在面对极度不平衡的数据集时,这一曲线提供了更为精确的性能评估。

GANomaly 与 AFPGAN 的 ROC 对比图和 PRC 对比图如图 9 和图 10 所示。

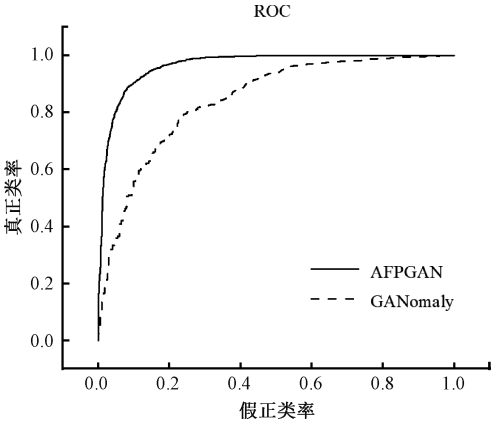


图 9 ROC 对比图

Fig. 9 ROC comparison diagram

基础模型和改进后的模型在重构损失和编码损失方面的表现,已经在图 11 和图 12 中进行了详细的展示和对比。据图 11 所示结果,可以看到 GANomaly 模型在训练过程中的重构损失曲线已逐渐趋于稳定,最终的收敛值为 0.058 9。而改进后的模型 AFPGAN 在相同训练条件下,其重构损失函数曲线呈现出更快的收敛趋势,最终的收敛值仅为 0.019 4。相比之下,AFPGAN 的重建损失

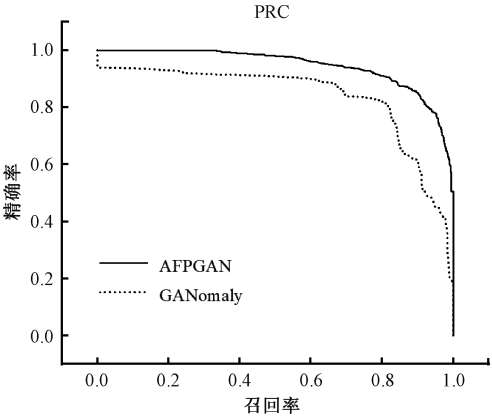


图 10 PRC 对比图

Fig. 10 PRC comparison diagram

降低了 0.039 5,这表明 AFPGAN 在重构能力方面得到了明显提升。同样,从图 12 的训练过程中可以看出,GANomaly 的编码损失函数曲线逐渐收敛于 0.039 6。改进后的模型 AFPGAN 最终收敛于 0.002 3。相比之下,AFPGAN 的编码损失降低了 0.037 3。这一结果表明,AFPGAN 在特征提取过程中不仅减少了模型在下采样过程中的关键信息损失,还提高了异常数据与正常数据之间的区分度。提高了模型的异常检测能力。

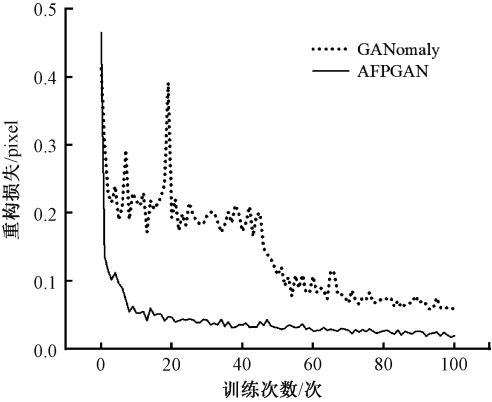


图 11 重构损失函数曲线对比图

Fig. 11 Comparison of reconstruction loss curves

2.5 对比实验

为了充分展示 AFPGAN 模型相较于其他模型的优越性,采用了相同的数据集对 GANomaly、F-AnoGAN、AnoGAN 以及 Skip-GANomaly 进行了训练和测试。具体结果如表 3 所示。

受试者工作特征曲线面积(area under the curve, AUC)是一个用于衡量分类器性能的统计指标,值越高表示模型的区分能力越强,即能更好地区分正常与异常。AFPGAN 在这一指标上得分为 0.964,显著高于 GANomaly、AnoGAN、F-AnoGAN 和 Skip-GANomaly 的分

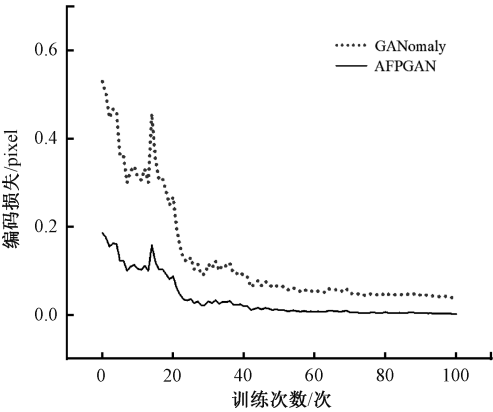


图 12 编码损失函数曲线对比图

Fig. 12 Comparison of Encoding Loss Curves

数,这意味着 AFPGAN 拥有更高的异常检测能力。平均精度 (average precision, AP) 值衡量的是模型在所有不同阈值下的精确度-召回率曲线下面积。在这个指标上,AFPGAN 的得分为 0.958,远超其他 4 个模型,表明了其不仅能够高效识别异常情况,而且在进行缺陷检测任务时维持了较高的精确度和召回率。

表 3 不同模型实验结果对比

Table 3 Comparison of experimental results of different models

模型	AUC	AP
GANomaly	0.858	0.852
AnoGAN	0.806	0.758
F-AnoGAN	0.682	0.704
Skip-GANomaly	0.814	0.808
AFPGAN	0.964	0.958

2.6 消融实验

为了验证在基础模型上引入的 NAM 注意力机制、流对齐模块和 PatchGAN 的有效性。进行一系列消融实验。所有实验均遵循先前提及的相同参数设置,以确保实验的一致性和准确性。实验结果如表 4 所示。

表 4 消融实验结果

Table 4 Results of ablation experiment

模型	AUC	AP
GANomaly	0.858	0.852
GANomaly+NAM	0.944	0.938
GANomaly+FAM	0.948	0.954
GANomaly+PatchGAN	0.918	0.933
GANomaly+NAM+FAM	0.953	0.941
GANomaly+NAM+PatchGAN	0.946	0.947
GANomaly+FAM+PatchGAN	0.956	0.952
GANomaly+NAM+FAM+PatchGAN	0.964	0.958

如消融实验数据所示,每个引入的模块在 AUC 和

AP 值上都有不小的提升,这些数据不仅凸显了每个模块的价值,也展现了它们在联合应用时能提供更为全面的性能提升。

3 结 论

AFPGAN 是以 GANomaly 为基础模型,其核心思路是在编码器部分引入 NAM 注意力机制,显著增强了模型的特征提取能力,从而使得模型能够更有效地捕捉到图像中的关键特征。在解码器部分引入了流对齐模块 FAM,增强了模型的重构能力。用 PatchGAN 代替原有的全局判别器,增强了判别器的鉴别能力。模型构建完成后,采用实验室自制数据集,对模型进行了训练和测试。通过一系列相关的实验,AFPGAN 在轮胎缺陷检测精度方面得到显著提升,AUC 和 AP 值均有较大提高。然而,实验中也发现较多流对齐模块导致模型在训练过程中显存占用率较高,并且在单独引入流对齐模块时,训练后期稳定性不足的现象。未来研究将集中于优化模型架构,调整流对模块的数量及其位置,以提高模型的鲁棒性,同时降低显存占用率。并观察能否进一步提升模型在缺陷检测中的性能。

参考文献

[1] 刘韵婷,刘鑫,高宇. 基于 FAMGAN 的轮胎 X 光图像缺陷检测[J]. 电子测量与仪器学报, 2023,37(12): 58-66.

LIU Y T, LIU X, GAO Y. Defect Detection of tire x-ray images based on FAMGAN [J]. Journal of Electronic Measurement and Instrumentation, 2023,37(12):58-66.

[2] 陈亮,吴攀,刘韵婷,等. 生成对抗网络 GAN 的发展与最新应用[J]. 电子测量与仪器学报, 2020,34(6): 70-78.

CHEN L, WU P, LIU Y T, et al. Development and latest applications of generative adversarial networks [J]. Journal of Electronic Measurement and Instrumentation, 2020,34(6):70-78.

[3] 乐华钢,王文武,朱磊,等. 融合自注意力与缺陷凸显的缺陷检测方法[J]. 电子测量与仪器学报, 2023, 37(9):85-92.

LE H G, WANG W W, ZHU L, et al. Defect detection method integrating self-attention and defect enhancement [J]. Journal of Electronic Measurement and Instrumentation, 2023,37(9):85-92.

[4] 张玥,陈锡伟,陈梦丹,等. 基于对比学习生成对抗网络的无监督工业品表面异常检测[J]. 电子测量与仪器学报, 2023, 37(10): 193-201.

ZHANG Y, CHEN X W, CHEN M D, et al. Unsupervised

- industrial product surface defect detection based on contrastive learning generative adversarial networks [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 193-201.
- [5] 伊欣同,单亚峰. 基于改进 Faster R-CNN 的光伏电池内部缺陷检测[J]. 电子测量与仪器学报, 2021, 35(1): 40-47.
- YI X T, SHAN Y F. Photovoltaic cell internal defect detection based on improved Faster R-CNN [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(1): 40-47.
- [6] 曾秀云,陆华才,吕禾丰. 基于改进 Faster R-CNN 的棉布包装缺陷检测的方法研究[J]. 电子测量与仪器学报, 2022, 36(4): 179-186.
- ZENG X Y, LU H C, LYU H F. Research on the method of cotton cloth packaging defect detection based on improved Faster R-CNN [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36 (4): 179-186.
- [7] 王标,周雅兰,王永红. 改进型 Faster R-CNN 网络在电子元件 LED 气泡缺陷检测中的应用[J]. 电子测量与仪器学报, 2021, 35(9): 136-143.
- WANG B, ZHOU Y L, WANG Y H. Application of the improved FasterR-CNN network in the detection of LED bubble defects in electronic components [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(9): 136-143.
- [8] 王宸,杨帅,周林,等. 基于自适应多尺度特征融合网络的金属齿轮端面缺陷检测方法研究[J]. 电子测量与仪器学报, 2023, 37(10): 153-163.
- WANG CH, YANG SH, ZHOU L, et al. Research on the defect detection method of metal gear end surface based on adaptive multi-scale feature fusion network [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 153-163.
- [9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in Neural Information Processing Systems, 2014, 27 (2): 2672-2680.
- [10] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery [C]. International conference on Information Processing in Medical Imaging, 2017: 146-157.
- [11] AKCAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Ganomaly: Semi-supervised anomaly detection via adversarial training[C]. Asian Conference on Computer Vision, 2019: 622-637.
- [12] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks [J]. Medical Image Analysis, 2019, 54: 30-44.
- [13] PERERA P, NALLAPATI R, XIANG B. Ocgan: One-class novelty detection using gans with constrained latent representations [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2019: 2898-2906.
- [14] HE L J, SHI N, MALIK K, et al. Unsupervised defect inspection algorithm based on cascaded GAN with edge repair feature fusion [J]. Applied Intelligence, 2022, 52(2): 2051-2069.
- [15] LI CH H, FU R ZH, LIU Y K. Algorithm for generating tire defect images based on RS-GAN [C]. International Conference on Neural Information, 2023: 388-399.
- [16] SAKURADA M, YAIRI T. Anomaly detection using autoencoders with nonlinear dimensionality reduction [C]. Machine Learning for Sensory Data Analysis, 2014: 4-11.
- [17] LIU Y CH, SHAO Z R, TENG Y Y, et al. NAM: Normalization-based attention module[J]. ArXiv preprint arXiv:2111. 12419, 2021.
- [18] LI X T, YOU AN SH, ZHU ZH, et al. Semantic flow for fast and accurate scene parsing [C]. European Conference Computer Vision, 2020: 775-793.
- [19] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]. IEEE Computer Vision and Pattern Recognition, 2017: 1125-1134.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [21] PARK J, WOO S, LEE J Y, et al. Bam: Bottleneck attention module[J]. ArXiv preprint arXiv:1807. 06514, 2018.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module [C]. European Conference on Computer Vision, 2018: 3-19.
- [23] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [J]. ArXiv preprint arXiv: 1502. 03167, 2015.
- [24] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks [J]. Advances in Neural Information Processing Systems, 2015, 28: 2010-2018.
- [25] DAI J F, QI H ZH, XIONG Y W, et al. Deformable convolutional networks [C]. IEEE International Conference on Computer Vision, 2017: 764-773.

作者简介



李洪奎, 2010 年毕业于沈阳工业大学电机与电器专业, 工学博士, 现为沈阳理工大学教授, 硕士研究生导师, 主要研究方向为电力系统及其自动化、人工智能技术。
E-mail: 1097601421@qq.com

Li Hongkui received his Ph. D. degree from the Department of Electrical Machines and Equipment at Shenyang University of Technology in 2010. He is now a professor and master's supervisor at Shenyang Ligong University. His main research interests include power systems and their automation, artificial intelligence technology.



陈浩, 2017 年于沈阳城市学院获得学士学位, 现为沈阳理工大学硕士研究生, 主要研究方向为异常检测。
E-mail: 1585436686@qq.com

Chen Hao received his B. Sc. degree from Shenyang City University in 2017. He is now a M. Sc. candidate at Shenyang University of Technology. His main research interest includes anomaly detection.



刘韵婷 (通信作者), 2015 年获东北大学博士学位, 现为沈阳理工大学副教授, 硕士生导师, 主要研究方向为无线传感器网络、人工智能技术和数据分析。
E-mail: liuyunting0224@163.com

Liu Yunting (Corresponding author) received her Ph. D. degree from Northeastern University in 2015. She is now an associate professor and master's supervisor of Shenyang Ligong University. Her main research interests include wireless sensor networks, artificial intelligence technology and data analysis.



张兴伟, 2022 年于沈阳工学院获得学士学位, 现为沈阳理工大学硕士研究生, 主要研究方向为异常检测。
E-mail: 1987792121@qq.com

Zhang Xingwei received his B. Sc. degree from Shenyang Institute of Technology in 2022. He is now a M. Sc. candidate at Shenyang University of Technology. His main research interest includes anomaly detection.



冯欣悦, 2023 年于山东工商学院获学士学位, 现为沈阳理工大学硕士研究生, 主要研究方向为异常检测。
E-mail: 1482323625@qq.com

Feng Xinyue received her B. Sc. degree from Shandong University of Commerce in 2023. She is now a M. Sc. candidate at Shenyang University of Technology. Her main research interest includes anomaly detection.