

基于双判别器的 GANomaly 异常检测方法研究



刘韵婷*, 谭明晓, 高宇, 戴佳霖

(沈阳理工大学 自动化与电气工程学院, 沈阳 110159)

摘要 在异常检测领域中, 生成对抗网络 (Generative Adversarial Nets, GAN) 和自编码器 (Autoencoder, AE) 近年来取得了较好的应用效果。然而, 现有的基于 GAN 的异常检测模型普遍存在重构能力差的问题。针对于此, 该文提出一种双判别器的 GANomaly 网络模型, 其中, 全局判别器用于提高图像的重构能力, 局部判别器用于提高在空间层次的编码能力。分别在 MvTec 数据集和自制轮胎 X 光图像数据集上对文中所提方法进行验证, 实验结果表明, 该方法能够有效提升模型的重构能力, 降低异常分数阈值, 提高异常检测的准确率。

关键词 异常检测; 深度学习; 双判别器; 生成对抗网络

中图分类号 TP391 文献标志码 A DOI 10.12178/1001-0548.2023157

Research on GANomaly Anomaly Detection Method Based on Double Discriminant

LIU Yunting*, TAN Mingxiao, GAO Yu, and DAI Jialin

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

Abstract In the field of anomaly detection, Generative Adversarial Nets (GAN) and Autoencoder (AE) have achieved better application results in recent years. However, the existing GAN-based anomaly detection models generally suffer from poor reconfiguration ability. To address this issue, this paper proposes a two-discriminator GANomaly network model, in which the global discriminator is used to improve the reconstruction ability of images and the local discriminator is used to improve the encoding ability at the spatial level. The proposed method is validated on the MvTec dataset and the homemade tire X-ray image dataset, respectively. The experimental results show that the proposed method can effectively improve the reconstruction ability of the model, reduce the anomaly score threshold, and improve the accuracy of anomaly detection.

Key words abnormality detection; deep learning; double discriminant; generating adversarial networks

异常检测^[1]原本是数据挖掘中的概念, 旨在查找出与预期不相符的数据, 通常这些数据被称为异常点, 也称畸变、例外、离群点等。异常检测在各个领域都有非常广泛地应用, 如金融^[2-3]、网络安全^[4-5]、工业检测^[6-8]、医疗诊断^[9-10]等。异常检测的经典传统算法有 PCA^[11]、KNN^[12]等, 在此基础上不断进行改进, 如将 One-Class SVM^[13]等单分类方法应用到异常检测, 虽取得了一定的效果, 但这些传统算法的可靠性和性能都有一定的缺陷。随着科技和计算机水平的不断进步, 人工智能、计算机视觉等得到了极大发展。深度学习在异常检测方面的

应用已经成为目前研究的热点, 根据相关研究表明, 基于深度学习的异常检测展现出来的效果已经超过了传统方法^[14]。因此, 基于深度学习的图像异常检测也越来越受到关注, 在一些复杂场景检测^[15]体现了重要作用。

异常检测的难点之一在于数据的高度不平衡, 通常情况下, 正常样本数据的比例远高于异常样本数据的比例, 这往往会导致实验结果不精确。利用无监督学习相关的异常检测模型是其中一种解决方法, 其基本思想是, 在训练阶段用正常的样本数据进行训练, 在测试阶段不符合正常样本数据特征的

收稿日期: 2023-06-06; 修回日期: 2023-09-22

基金项目: 辽宁省教育厅面上项目 (LJKMZ2022061); 辽宁省自然科学基金 (2022-KF-14-022)

作者简介: 刘韵婷, 博士, 副教授, 主要从事深度学习、数据分析及无线传感器网络等方向的研究。

*通信作者 E-mail: 71019976@qq.com

就被判为异常。

目前基于 GAN^[16] 网络的异常检测方法已经取得了非常大的突破。文献 [17] 提出了 AnoGAN 模型, 该方法在训练时只用正常样本, 在测试时, 由于 GAN 只学习了正常样本的数据分布, 因此, 重构图像在保留原有图像特征的基础上去除缺陷的部分, 再利用重构图像与原图像的残差确定缺陷的位置。文献 [18] 利用生成对抗网络和最小似然正则化来实现领域异常检测, 通过增加生成器生成的异常样本, 来防止异常样本收敛到正常数据分布。文献 [19] 提出无须异常样本实现异常检测的 GANomaly 模型, 该模型通过编码器对输入的图像进行编码得到输入图像的低维向量, 再通过解码器和编码器得到重构图像的潜在向量, 二者的差值超出异常阈值分数则被判定为异常。但在实践的过程中发现 GANomaly 网络模型重构图像的能力存在不足。在此基础上进行改进的 skip-GANomaly^[20] 模型, 去掉了再编码的步骤, 而替代的是在自编码器中加入跳跃连接, 将浅层的特征信息与深层的特征信息进行融合。改良后的 skip-GANomaly 模型在某些方面取得了很好的效果, 但受目标特征的尺寸大小影响较大。还有很多关于 GAN 的异常检测模型, 如 BiGAN^[21]、深度卷积生成对抗网络 (Deep Convolutional Generative Adversarial Networks, DCGAN)^[22]、自编码器对抗网络 (Autoencoder Adversarial Network)^[23]、基于 GAN 的异常检测框架-对抗性双自动编码器^[24] 等, 这些模型普遍存在重构图像能力不足的问题。

本文针对 GANomaly 模型对图像重构能力不强和异常分数阈值不够小的问题, 在 GANomaly 模型的基础上添加了一个局部判别器, 当输入正常图像进行训练时, 全局判别器使原图像和重构图像的重构误差尽可能小, 使模型具有更好的图像重构能力, 局部判别器使得原图像经过编码器产生的低维向量和重构图像经过编码器产生的潜在向量误差尽可能小, 使模型的异常阈值分数尽可能的小, 全局判别器和局部判别器同时进行对抗学习, 使得该模型的架构更加完善。最终通过 MvTec 数据集和自制的轮胎 X 光图像数据集的相关实验证明了该网络能有效提升异常检测精度。

1 双判别器 GANomaly 网络模型

1.1 GANomaly

GANomaly 是基于 GAN 改进的模型。主要采

用编码器-解码器-编码器结构, 可以将输入的图像通过第一个编码器映射成一个低维向量, 然后通过解码器将低维的向量重构为图像, 最后再通过第二个编码器得到重构图像的潜在向量。通过模型的训练, 可以进一步地缩小低维向量与潜在向量之间的差别, 从而学习并得到输入图像的数据分布, 判别器利用从该学习数据分布获取的较大的距离度量去推断该分布中存在的异常值^[25]。

该网络一共由 3 部分组成, 第一部分是自编码器, 作为该网络的生成器。网络首先读取图像 x 到编码器 $1(G_E(x))$, 得到低维向量 Z , 即 $Z = G_E(x)$, 然后通过解码器得到重构图像 x' , 即 $x' = G_D(Z)$ 。第二部分是再编码器, 结构与编码器 1 一样, 主要作用是得到重构图像 x' 的潜在向量 Z' , 即 $Z' = G_E(x')$ 。第三部分是判别器网络 D , 主要作用是区分自身输入的图像是来自真实图像 x 还是重构的图像 x' 。

测试时, 用训练好的编码器-解码器-编码器模型。由于训练时采用正常的图像, 当输入异常图像时, 整个结构会尽量重构成正常图像, 使得第一个编码器产生的低维向量与再编码器产生的潜在向量的差距变大, 超过一定的异常阈值分数就被判定为异常。

1.2 DDGANomaly

为增强 GANomaly 模型的重构能力, 降低异常阈值分数, 提高检测精度。本文在已有的 GANomaly 模型的基础上引入双对抗机制, 即两个判别器具有不同功能。局部判别器使输入图像经过自编码器中的编码器生成的低维向量与重构图像再压缩生成的潜在向量更加接近, 全局判别器可以进一步减小真实图像和重构图像的重构误差。相关实验证明, 双对抗机制的引入, 可以减小重构误差, 降低异常阈值分数, 有效提高模型异常检测的精度。

双判别器 GANomaly (Dual Discriminator GANomaly, DDGANomaly) 的模型结构如图 1 所示。改进后的模型由原来的 3 个部分增加到 4 个部分, 分别为自编码器、再编码器、全局判别器以及局部判别器。其中自编码器、再编码器、全局判别器 3 部分的模型结构与传统的 GANomaly 模型结构相同。增加的局部判别器在模型结构上与全局生成器相比去除了编码器。

与基础 GANomaly 模型相同, DDGANomaly 模型的优化目标是判别器和生成器相互对抗的问题, 与基础 GANomaly 模型相比不同之处在于 DDGANomaly 中存在两个对抗过程。

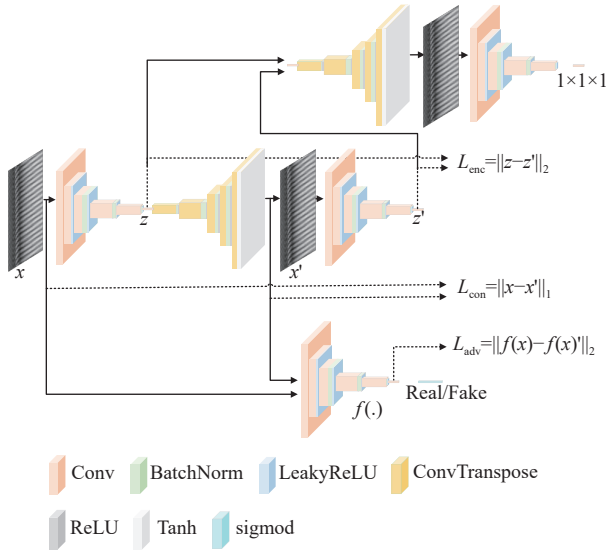


图 1 DDGANomaly 结构图

1) 自编码器和再编码器组成的全局生成器与全局判别器进行对抗, 使输入的真实图像和重构图像之间的差值最小化。

2) 局部判别器与再编码器构成的生成器进行对抗。使得自编码器中编码器产生的低维向量与再编码器产生的潜在向量的差值尽可能小。

1.3 损失函数

1.3.1 全局生成器和全局判别器损失函数

1) 全局生成器损失函数

全局生成器的损失函数包括 3 个, 分别为特征匹配性损失 L_{adv} 、重构误差损失 L_{con} 和编码损失 L_{enc} , 判别器的损失函数就采用传统的生成对抗网络的交叉熵损失函数。

2) 特征匹配损失

L_{adv} 表示特征匹配的误差, 通过计算输入图像和重构图像在判别器中间层的特征表示的欧式距离, 使得 x 和 x' 更接近, 让生成器生成的图像更逼真, 同时进一步优化判别器:

$$L_{adv} = \|f(x) - f(G(x))\|_2 \quad (1)$$

3) 重构误差损失

L_{con} 表示重构误差损失, 通过该函数不断减小重构图像与原图像的差异:

$$L_{con} = \|x - x'\|_1 \quad (2)$$

4) 编码损失

L_{enc} 表示输入图像的低维向量和重构图像的潜在向量的差异, 即:

$$L_{enc} = \|Z - Z'\|_2 \quad (3)$$

5) 其目标函数为:

$$L = w_{adv}L_{adv} + w_{enc}L_{enc} + w_{con}L_{con} \quad (4)$$

式中, w_{adv} 、 w_{enc} 、 w_{con} 为加权参数。

6) 全局判别器损失

判别器损失采用的就是原始 GAN 的交叉熵损失函数, 公式如下:

$$\min_G \max_D F(D, G) = E_{x \sim P_{data}(x)} \log D(x) + E_{z \sim P_Z(z)} \log(1 - \log D(G(z))) \quad (5)$$

7) 异常阈值分数

利用自编码器中编码器产生的低维向量与重构图像再压缩产生的潜在向量差作为异常阈值分数, 其差值越小代表模型训练的效果越好。

异常分数定义如下:

$$A(a) = \|G_E(a) - E(G(a))\|_1 \quad (6)$$

由模型的训练得到一个异常阈值分数 φ , 测试图像得到的异常分数 $A(a) > \varphi$, 表示该图像为异常。

1.3.2 局部生成器和局部判别器损失函数

局部判别器的目的主要是缩小输入图像经过自编码器产生的低维向量和再编码器产生的重构图像的潜在向量的差异。局部判别器和局部生成器的目标损失函数采用 Wasserstein^[26] 距离损失函数, 避免出现模式崩溃^[27] 等问题。

具体进行以下 4 点改进。

1) 局部判别器最后一层去掉激活函数

Wasserstein 距离损失函数, 对于局部判别器的输出要求是实际的输出值。基本的 GAN 模型中的判别器最后一层一般是 sigmoid 激活函数。为得到局部判别器的实际输出值, 去掉 sigmoid 激活函数。

2) 使用基于动量的优化算法 RMSProp

采用 Adam 优化算法可能会引起训练的不稳定, 因此采用 RMSProp 优化算法。

3) 控制参数的范围

每次更新判别器的参数后将其映射到一个固定常数 ω 的范围内。

4) 替换损失函数

将真实图像与生成器图像输入局部判别器中, 用二者输出实际的差值代替 BCELoss 损失。对局部判别器优化, 二者的差值越大越好。对于局部生成器的优化, 局部判别器对局部生成器生成的图像产生的实际输出值越大越好。

局部生成器损失函数和局部判别器损失函数公

式如式(7)和式(8)所示:

$$\min_G = -F_{x \sim P_{\text{data}(x)}}[D_w(x)] \quad (7)$$

$$L(D) = -F_{x \sim P_{\text{data}(x)}}[D_w(x)] + E_{z \sim P_{Z(Z)}}[D_w(x)] \quad (8)$$

1.4 训练方式及测试

DDGANomaly模型训练算法的细节如算法1所示,此算法共有3个训练过程,分别是全局生成器优化以及两组判别器的对抗过程。

算法1 双判别器GANomaly的训练过程

输入: 原始样本集 $X = \{x_i\}_{i=1}^M$

全局和局部判别器的训练次数为 T

repeat

for $i=1$ to M do

$X_i \leftarrow i$ th batchsize

for $T=1$ to T do

通过对抗性损失、编码器损失、上下文损失更新自编码器和编码器部分:

$$L = w_{\text{adv}}L_{\text{adv}} + w_{\text{enc}}L_{\text{enc}} + w_{\text{con}}L_{\text{con}}$$

end for

for $T=1$ to T do

通过交叉熵损失更新全局判别器:

$$\min_G \max_D F(D, G) = F_{x \sim P_{\text{data}(x)}} \log D(x) +$$

$$E_{z \sim P_{Z(Z)}} \log(1 - \log D(G(Z)))$$

end for

for $T=1$ to T do

通过 Wasserstein 距离损失函数更新局部判别器:

$$L(D) = -F_{x \sim P_{\text{data}(x)}}[D_w(x)] + E_{z \sim P_{Z(Z)}}[D_w(x)]$$

end for

end for

until 达到纳什均衡

训练完成后,全局判别器无法辨别输入是生成器生成的重构图像还是真实图像,局部判别器也无法分辨出输入是来自编码器中的低维向量还是重构图像经过再编码器生成的潜在向量。

测试阶段利用训练好的自编码器和再编码器组成的测试结构对待检测的样本数据进行检测。在模型的训练阶段全部选取正常样本数据,因此训练好的模型对正常样本数据具有较强的重构能力,对异常样本数据重构能力较差。将待检测的样本数据输入测试结构中,将编码损失与训练好的异常阈值分数进行比较,如式(6)所示,如果编码损失的值超过异常阈值分数就被判别为异常样本数据,否则被

判别为正常样本数据。

2 实验结果与分析

2.1 实验数据集制作

本文准备了两个数据集,分别是 MvTec^[28]数据集和自制的轮胎 X 光图像数据集。

2.1.1 MvTec 数据集

MvTec 数据集一共有 15 个不同种类的缺陷,其中用于训练与验证的图片共有 3 629 张,用于测试的有 1 725 张,用于训练的只有正常的图片。测试集中包含 73 种不同的缺陷,如物体表面缺陷、结构缺陷或者由于缺少某些物体组成部件而表现出来的缺陷。

所有图像的尺寸在 $700 \times 700 \sim 1\,024 \times 1\,024$ 之间,其中网格、拉链和螺丝图像为灰度单通道图像。数据集为所有的缺陷图像都给出了像素级标注区域,涵盖了接近 1 900 个标注区域。网格的部分图片如图 2 所示。

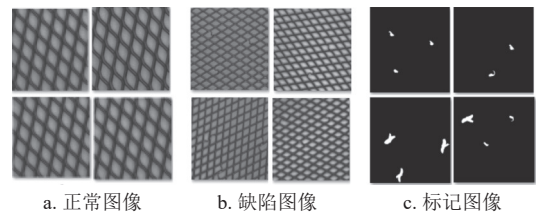


图2 网格数据集

2.1.2 自制轮胎数据集

1) 训练数据集

训练数据集由 13 200 张无缺陷的轮胎 X 光图像构成。具体选用子午线轮胎 X 光图像,并经过轮胎不同位置的 X 光图像对比,最后确定选用结构和背景相对简单的胎侧部分的 X 光图像,如图 3 所示。

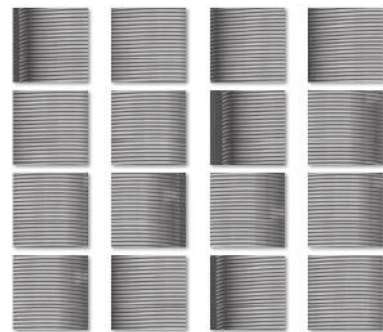


图3 轮胎 X 光图像训练集正常轮胎样本

2) 测试数据集

测试数据集由 3 部分构成:正常的轮胎 X 光图像、含有气泡的轮胎 X 光图像、含有杂质的轮

胎 X 光图像, 分别有 500 张图片, 共计 1 500 张。含有气泡和杂质的轮胎 X 光图像如图 4 所示。

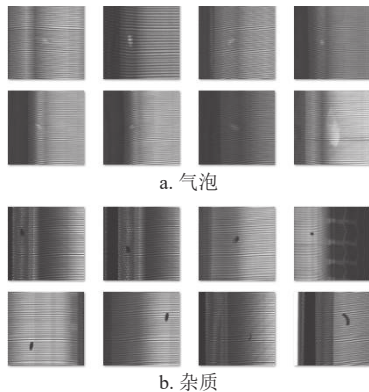


图 4 轮胎 X 光图像测试集缺陷样本

2.2 实验相关参数

1) 参数设置

优化器: Adam; lr=0.000 1; $b_1=0.5$; $b_2=0.999$; 隐变量维度 $Z_1=100$; batchsize=132; epoch=100/50。

2) 实验平台配置

计算机的硬件配置为 Intel-i5、Nvidia GTX1050 GPU、12G 内存。编程环境采用 windows 操作系统, Python3.6 及 Pytorch 深度学习开发工具。

2.3 实验结果分析

本次实验的模型性能的评价方法采用受试者工作特征 (Re-ceiver Operating Characteristic, ROC) 曲

线。ROC 曲线的横轴代表负阳率, 表示异常数据被错误地判别为正常的比例; 纵轴代表正阳率, 表示正常数据被正确地判别为正常的比例。ROC 的曲线越贴近坐标轴的左上角代表检测的效果越佳。ROC 曲线与横轴围成的面积代表 AUC 的值, 该值越大, 同样代表检测的效果越优秀。单一的 ROC 曲线对于实验结果的验证缺乏单一性, 采用 AUC 可以更好验证实验的结果。以轮胎 X 光图像数据集为例, 原 GANomaly 与改进后的模型 DDGANomaly 的 AUC 对比如图 5 所示。

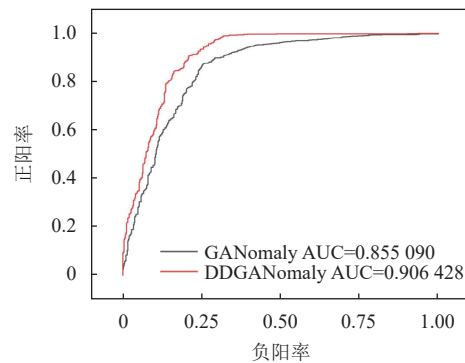


图 5 轮胎 X 光图像缺陷检测 ROC 曲线图对比图

2.3.1 MvTec 数据集

为了论证所提的方法在 MvTec 数据集上的有效性, 选取先进的无监督异常检测模型进行对比实验, 实验结果如表 1 所示。

表 1 MvTec 数据集 AUC 值实验结果对比表

种类	模型				
	Autoencoder	AnoGAN	f-AnoGAN	GANomaly	DDGANomaly
bottle	0.457	0.491	0.482	0.635	0.711
Cable	0.508	0.526	0.614	0.648	0.649
capsule	0.589	0.412	0.534	0.505	0.548
carpet	0.721	0.472	0.532	0.437	0.533
grid	0.718	0.610	0.800	0.856	0.891
hazelnut	0.814	0.649	0.622	0.823	0.833
leather	0.899	0.342	0.77	0.579	0.545
metal_nut	0.532	0.382	0.396	0.466	0.520
pill	0.348	0.416	0.356	0.458	0.564
screw	0.328	0.721	0.879	0.998	0.997
tile	0.506	0.667	0.723	0.457	0.534
toothbrush	0.486	0.539	0.497	0.798	0.864
transistor	0.686	0.561	0.512	0.792	0.823
wood	0.515	0.578	0.901	0.821	0.847
zipper	0.346	0.437	0.378	0.642	0.664
均值	0.560	0.520	0.599	0.661	0.710

从表1可以看出,改进后的模型 DDGANomaly,对 MvTec 数据集的 15 种不同类型的缺陷其 AUC 的均值为 0.71,相比于原 GANomaly 模型的 AUC 均值提升了 0.049,增加局部判别器之后,DDGANomaly 模型的重构能力更强,异常阈值分数降低,检测结果更加精确。并且相比于另外 3 种无监督异常检测模型 Autocoder、AnoGAN 和 f-AnoGAN 的 AUC 均值分别提升了 0.150、0.190 和 0.111,这是由于 Autocoder、AnoGAN 和 f-AnoGAN 这 3 种异常检测模型都是基于图像的检测判别,而缺陷的种类丰富多样并且一些图片的背景非常复杂,在重构图像时会不准确,导致异常检测的精确度不高。实验证明了该方法是有效的。不同模型,同一种类缺陷的 AUC 值进行对比,Autoencoder 模型的 capsule、carprt、leather 和 metal_nut 的 AUC 值略高于改进后的模型。除此之外,AnoGAN 模型的 tile、f-AnoGAN 模型的 leather、tile、wood,以及 GANomaly 模型的 leather 和 screw 的 AUC 值,都略高于改进后的模型 DDGANomaly。这是由于不同缺陷种类的图像背景复杂度不同,不同模型面对同一种缺陷类型的特征提取能力有差别,导致个别模型对某一种缺陷类型的检测精度较高。除此之外,不同模型对应的缺陷种类的 AUC 值都低于改进后的 DDGANomaly 模型。通过实验论证,改进后的模型 DDGANomaly 对于异常检测具有较强的普适性。

2.3.2 轮胎 X 光图像数据集

1) 训练损失

原始 GANomaly 模型与改进后的 DDGANomaly 模型的编码损失函数和重构误差损失函数分别如图6和图7所示。

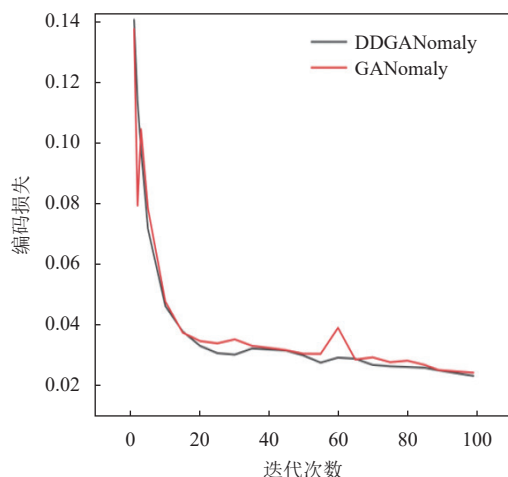


图6 编码损失函数曲线图

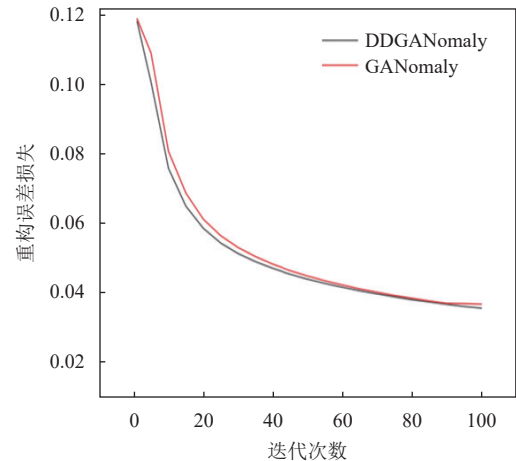


图7 重构误差损失函数曲线图

由图6的编码器损失函数曲线可知,GANomaly 模型的编码损失的收敛值为 0.024 5,改进后的模型 DDGANomaly 收敛值为 0.023 4,编码器的损失下降了 0.001 1,降低了异常分数阈值。同样由图7训练的重构误差损失函数曲线图可知,GANomaly 的重构误差损失的收敛值为 0.036 6,改进后的模型 DDGANomaly 的收敛值为 0.035 4,重构误差损失下降了 0.001 2,增强了模型的重构能力。

2) 检测结果

同样所提的方法在自制的轮胎 X 光图像数据集上对各类缺陷的识别结果,如表2所示。

表2 轮胎 X 光图像 AUC 值实验结果对比

模型	AUC	准确率
Autoencoder	0.597	0.602
AnoGAN	0.710	0.669
f-AnoGAN	0.604	0.624
GANomaly	0.855	0.829
DDGANomaly	0.906	0.888

由表2可以看出,在轮胎 X 光图像数据集上的检测效果与 MvTec 数据集上的检测效果一致,改进的模型 DDGANomaly 以 0.906 的 AUC 值领先其他无监督异常检测模型,相比 GANomaly 模型、Autoencoder 模型、AnoGAN 模型和 f-AnoGAN 模型的 AUC 值分别提高了 0.051、0.309、0.196 和 0.302。

其中,在 MvTec 数据集上表现较好的两个无监督异常检测模型 Autocoder 和 f-AnoGAN,在自制的轮胎 X 光图像数据集上表现不如 AnoGAN 模型,这是由于轮胎 X 光图像的背景十分复杂,图像背景和缺陷区域很难分别, Autocoder 和 f-AnoGAN 在模型结构上都存在编码器-解码器的结构,并且

最后都是通过将模型重构的生成图像与输入的原始图像进行比较来辨别是否存在异常, 而这样的结构在重构一些较为复杂背景的图像时, 重构能力较差, 导致检测精度下降。

3) 消融实验研究

通过消融实验来验证本文所提出的在 GANomaly 模型中增加局部判别器和引入 Wasserstein 距离损失函数的有效性。相同实验条件下, 评估不同模块对本文模型性能的影响, 消融实验选择 GANomaly 作为基准模型, 在处理后的轮胎 X 光图像数据集上进行实验, 相关的实验参数与上述一致, 实验结果如表 3 所示。

GANomaly+D 表示在原有 GANomaly 基础上添加局部判别器, 增加局部判别器之后, 将自编码器生成的低维向量和再编码器生成的潜在向量输入局部判别器中, 二者经过不断地对抗, 缩小低维向量和潜在向量的误差, 减小编码损失, 进而提高模型的重构能力, 降低异常阈值分数, 在检测结果上, 检测的精度上升了 4%。GANomaly+D+w 表示在原有的 GANomaly 模型的基础上添加局部判别器之后, 再添加 Wasserstein 距离损失函数, Wasserstein 距离损失函数的引入, 避免在训练模型时出现梯度消失的情况, 从而提高模型训练的稳定性, 可以使训练出来的结果更加精确, 同样在检测结果上, 检测的精度提升了 1.1%, 证明了对其改进是有效的。

表 3 消融实验数据表

模型	AUC	准确率
GANomaly	0.855	0.829
GANomaly+D	0.895	0.882
GANomaly+D+w	0.906	0.888

3 结束语

本文提出的 DDGANomaly 模型, 在原始 GANomaly 模型的基础上添加局部的判别器, 形成双判别器机制, 有效解决了由于自编码器的存在而导致模型图像重构能力差以及原始 GANomaly 模型中异常阈值分数过高的问题。通过相关的实验验证, 改进后的模型 DDGANomaly 能够提升异常检测的准确率。

参考文献

[1] 王军. 基于多尺度特征预测的异常事件检测[J]. 电子科技大学学报, 2022, 51(4): 586-591.

- WANG J. Abnormal event detection based on multi-scale features prediction[J]. *Journal of University of Electronic Science and Technology of China*, 2022, 51(4): 586-591.
- [2] 钟增胜, 朱纯瑶, 杨逸飞, 等. 数字货币交易所洗钱行为检测[J]. 湖南大学学报(自然科学版), 2022, 49(10): 119-129.
- ZHONG Z S, ZHU C Y, YANG Y F, et al. Money laundering detection for cryptocurrency transactions[J]. *Journal of Hunan University (Natural Sciences)*, 2022, 49(10): 119-129.
- [3] JENNY D, DMITRIY Y, DMITRIY B, et al. Detection and analysis of atypical stock transactions with possible misuse of insider information and market manipulation: Methods and models[J]. *Procedia Computer Science*, 2022, 213: 165-174.
- [4] SHEERAZ A, ALI K Z, MUHAMMAD M S, et al. Effective and efficient DDoS attack detection using deep learning algorithm, multi-layer perceptron[J]. *Future Internet*, 2023, 15(2): 76.
- [5] MURRAY A, RAWAT D B. Towards botnet hazard analysis with generative adversarial networks for threat detection[C]//Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III. [S.l.]: SPIE, 2021: 201-212.
- [6] QIN C, WU R, HUANG G, et al. A novel LSTM-autoencoder and enhanced transformer-based detection method for shield machine cutterhead clogging[J]. *Science China Technological Sciences*, 2023, 66(2): 512-527.
- [7] MIKOŁAJ O, JUYOUL K. Detection of numerical power shift anomalies in burnup modeling of a PWR reactor[J]. *Sustainability*, 2023, 15(4): 3373.
- [8] EZZAT D, HASSANIEN A E, DARWISH A, et al. Multi-objective hybrid artificial intelligence approach for fault diagnosis of aerospace systems[J]. *IEEE Access*, 2021, 9: 41717-41730.
- [9] NAWAZ M, NAZIR T, BAILI J, et al. CXray-EffDet: Chest disease detection and classification from X-ray images using the EfficientDet model[J]. *Diagnostics*, 2023, 13(2): 248.
- [10] WEKKING D, PORCU M, DE SILVA P, et al. Breast MRI: Clinical indications, recommendations, and future applications in breast cancer diagnosis[J]. *Current Oncology Reports*, 2023, 25(4): 257-267.
- [11] 冯安然, 王旭仁, 汪秋云, 等. 基于 PCA 和随机树的数据库异常访问检测[J]. 计算机科学, 2020, 47(9): 94-98.
- FENG A R, WANG X R, WANG Q Y, et al. Database anomaly access detection based on principal component analysis and random tree[J]. *Computer Science*, 2020, 47(9): 94-98.
- [12] 陈彬杰, 魏福山, 顾纯祥. 基于 KNN 的具有隐私保护功能的区块链异常交易检测[J]. 信息安全, 2022, 22(3): 78-84.
- CHEN B J, WEI F S, GU C X. Blockchain abnormal transaction detection with privacy-preserving based on KNN[J]. *Netinfo Security*, 2022, 22(3): 78-84.
- [13] ZHOU L C, LU T, HU X B. Detecting web application injection attacks using one-class SVM[C]//Proceedings of

- the IEEE 5th International Conference on Computer and Communication Engineering Technology. New York: IEEE, 2022: 275-279.
- [14] JAVAID A, NIYAZ Q, SUN W Q, et al. A deep learning approach for network intrusion detection system[C]//Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS). New York: ACM, 2016: 21-26.
- [15] TAN Q M, SUN Y K, XIN P F, et al. An approach on visual detecting multi-targets in the unstructured and complex scenes based on RGB-D images[C]//Proceedings of the 5th International Conference on Automation, Control and Robotics Engineering. New York: IEEE, 2020: 591-596.
- [16] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[EB/OL]. [2023-09-21]. <http://arxiv.org/abs/1406.2661v1>.
- [17] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery[EB/OL]. [2023-09-21]. <http://arxiv.org/abs/1703.05921v1>.
- [18] WANG C, ZHANG Y M, LIU C L. Anomaly detection via minimum likelihood generative adversarial networks [C]//Proceedings of the 24th International Conference on Pattern Recognition. New York: IEEE, 2018: 1121-1126.
- [19] AKÇAY S, ATAPOUR-ABARGHOUEI A, BRECKON T. GANomaly: Semi-supervised anomaly detection via adversarial training[C]//14th Asian Conference on Computer Vision. Perth: Springer, 2019: 622-637.
- [20] AKÇAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Skip-GANomaly: Skip connected and adversarially trained encoder-decoder anomaly detection[C]//Proceedings of the International Joint Conference on Neural Networks. New York: IEEE, 2019: 1-8.
- [21] KAPLAN M O, ALPTEKIN S E. An improved BiGAN based approach for anomaly detection[J]. *Procedia Computer Science*, 2020, 176: 185-194.
- [22] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. [2023-09-21]. <https://arxiv.org/abs/1511.06434>.
- [23] MAKHZANI A, SHLENS J, JAITLY N, et al. Adversarial autoencoders[EB/OL]. [2023-09-22]. <http://arxiv.org/pdf/1511.05644v1>.
- [24] VU H S, UETA D, HASHIMOTO K, et al. Anomaly detection with adversarial dual autoencoders[EB/OL]. [2023-09-21]. <http://arxiv.org/abs/1902.06924v1>.
- [25] 王齐, 陈功, 胡文昕, 等. 使用 GANomaly 网络的面瘫识别应用研究[J]. *软件工程*, 2022, 25(3): 29-33.
WANG Q, CHEN G, HU W X, et al. Application research of facial paralysis recognition based on GANomaly network[J]. *Software Engineering*, 2022, 25(3): 29-33.
- [26] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//Proceedings of the 34th International Conference on Machine Learning. New York: ACM, 2017: 214-223.
- [27] NGUYEN T D, LE T, VU H, et al. Dual discriminator generative adversarial nets[EB/OL]. [2023-09-21]. <http://arxiv.org/abs/1709.03831v1>.
- [28] 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. New York: IEEE, 2019: 9584-9592.

编辑 张莉