

DOI: 10.13382/j.jemi.B2306808

# 基于 FAMGAN 的轮胎 X 光图像缺陷检测\*

刘韵婷 刘鑫 高宇

(沈阳理工大学自动化与电气工程学院 沈阳 110159)

**摘要:**针对气泡缺陷特征和图像背景像素差异较小、检测困难的问题,以 Skip-GANomaly 为基础框架,提出了融合注意力机制生成对抗网络(FAMGAN),首先,生成器中编码器和解码器之间的跳连层由注意力特征融合模块(AFF)和注意力机制模块(CBAM)构成,提高了对目标特征的关注、减少了图像特征丢失;然后,在判别器中加入联合上采样模块(JPU),提高了模型检测图像缺陷的速度。最后,将本文提出的 FAMGAN 网络与近几年经典的生成对抗网络在自制的轮胎缺陷数据集上进行训练、测试和评估。实验结果表明,本文提出的网络对轮胎气泡缺陷检测的精度达到 0.837,相比于 Skip-GANomaly 网络提高了近 30%。

**关键词:**生成对抗网络;CBAM;深度学习;AFF;轮胎图像缺陷检测;JPU

**中图分类号:** TP274 **文献标识码:** A **国家标准学科分类代码:** 510.4

## Defect detection of tire X-ray images based on FAMGAN

Liu Yunting Liu Xin Gao Yu

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

**Abstract:** In response to the problem of small differences in blister defect features and background pixels in tire defect images, as well as difficulty in detection, Skip-GANomaly is adopted as the basic framework to propose the fusion attention mechanism generative adversarial network (FAMGAN). Firstly, the skip layer between the encoder and decoder in the generator consists of an attention feature fusion (AFF) module and a convolutional block attention module (CBAM) module, which improves the focus on target features and reduces image feature loss. Then, a joint pyramid upsampling (JPU) module was added to the discriminator to improve the speed of the model in detecting image defects. Finally, the FAMGAN network proposed in this article will be trained, tested, and evaluated on a self-made tire defect dataset with classic generative adversarial networks in recent years. The experimental results show that the proposed network achieves an accuracy of 0.837 for tire blister defect detection, which is nearly 30 percentage points higher than the Skip GANomaly network.

**Keywords:** generative adversarial network; CBAM; deep learning; AFF; tire image defect detection; JPU

## 0 引言

轮胎是汽车和地面唯一有接触的部分,为避免轮胎缺陷引起的交通事故,需要严格把控出厂轮胎的质量,轮胎缺陷检测作为保证轮胎质量的方法之一。目前,轮胎缺陷检测主要采用人工检测法,依靠人工经验对获取的轮胎图像进行观察和判断。虽然此方法使用简单,但是

实际应用过程中主观性较强、资源消耗较大,不利于大规模的工业生产。为了有效解决传统人工缺陷检测法的问题,基于深度学习的图像缺陷检测方法应运而生。

随着深度学习的发展,基于深度学习图像缺陷检测已经取得了很大的成功<sup>[1-4]</sup>。2014年,Mirza等<sup>[5]</sup>首次提出了生成对抗网络(generative adversarial network, GAN),GAN是一种典型的无监督学习方法。GAN由生成器和判别器组成,生成器用于重建原始输入图像,判别

收稿日期: 2023-08-07 Received Date: 2023-08-07

\* 基金项目:辽宁省自然科学基金(2022-KF-14-02)、辽宁省教育厅面上项目(LJKMZ20220617)资助

器用于区分重建图像和原始输入图像,并将信息反馈给生成器,迫使生成器重建和输入图像更加相似的图像,生成器和判别器在对抗中相互学习,当判别器无法区分重建图像和原始输入图像时,训练结束。Schlegl 等<sup>[6]</sup>首次提出使用生成对抗网络(anomaly detection with generative adversarial networks, AnoGAN)进行图像缺陷检测,AnoGAN 使用卷积神经网络获得原始输入图像的特征分布,并将生成图像和原始图像输入判别器进行分类,最后,通过计算原始图像和生成图像之间的残差确定异常值。但 AnoGAN 的计算量非常大,训练速度比较慢。为了解决 AnoGAN 训练速度慢的问题,2019 年,Schlegle 等<sup>[7]</sup>提出了 F-AnoGAN,在 AnoGAN 的基础上又加入了一个自编码器,与 AnoGAN 相比,F-AnoGAN 可以将图像快速映射到潜在空间的某一点。GANomaly<sup>[8]</sup>采用了编码器-解码器-编码器模型结构,首先通过下采样得到样本的潜在特征向量,并利用解码器上采样进行图像重建,然后,利用第 2 个编码器对生成图像下采样得到生成图像的潜在特征向量,并计算生成图像和输入图像的潜在特征向量差,从而提高模型图像重建能力。但 GANomaly 重建复杂图像的能力并不理想。2020 年,Tang 等<sup>[9]</sup>提出了 DAGAN,其生成器和判别器都采用了自编码器,提高了模型的训练稳定性,仅使用少量的图像进行训练就能得到较好的训练结果,但使用大量的图像进行训练,得到的结果并不理想。2021 年,Wang 等<sup>[10]</sup>等在 DCGAN 的基础上,提出了 MAGAD,在生成器的编码器和解码器之间加入了 memAE<sup>[11]</sup>,增强了模型记忆图像特征的能力,但使用分辨率较高的图像训练 MAGAD 得到的训练效果依然不理想。Niu 等<sup>[12]</sup>提出了一个扩充换向器表面缺陷的 GAN (surface defect-generation adversarial network, SDGAN),当光照不均以及光线较暗时,SDGAN 具有较好的鲁棒性。目前,生成对抗网络技术已经应用于制造业<sup>[13]</sup>、视频监控<sup>[14]</sup>和面部表情合成<sup>[15]</sup>等许多领域。

Mei 等<sup>[16]</sup>构建了一种基于高斯金字塔的 CDAE 架构,高斯金字塔用于实现不同分辨率的图像重建。

受 Resnet<sup>[17]</sup>的启发,Ronneberger 等<sup>[18]</sup>提出了 U-net。2019 年,Akay 等<sup>[19]</sup>提出了 skip-GANomaly,在生成器的编码器和解码器之间加入了 skip-connection,并使用跳层连接将转置卷积(ConTranspose)和卷积(Conv)的输出特征向量进行融合,该方法可以减少图像特征丢失,提高了输入图像和生成图像的相似度。此外,该方法具有较好的图像重建能力而且训练比较稳定。

现阶段生成对抗网络的生成器大多采用自编码器(auto encoder, AE)模型结构。2020 年,Liu<sup>[20]</sup>等提出了二维主成分分析卷积自编码器(two-dimensional principal

component analysis-based convolutional autoencoder, PCACAE),PCACAE 通过提取非线性图像特征,可以获得更多具有价值的特征信息。

自编码器可以提取图像特征进行图像重建。但自编码器很难同时关注图像的目标特征和全局特征,如果图像缺陷和图像背景相似度较高,会导致模型很难分辨出图像是否具有缺陷。由于这个原因,目前提出的 GAN 模型检测气泡缺陷,得到的结果并不理想。

为了解决这个问题同时提高模型学习图像特征的能力。受 skip-GANomaly 以及 U-net 的启发,本文的基本框架采用了去掉跳层连接的 Skip-GANomaly 网络,提出了融合注意力机制的生成对抗网络 FAMGAN (fusion of attention mechanism generative adversarial network)。使用注意力特征融合模块(attention feature fusion, AFF)<sup>[21]</sup>、注意力机制(convolutional block attention module, CBAM)<sup>[22]</sup>以及联合上采样(joint pyramid upsampling, JPU)<sup>[23]</sup>对基本框架的结构进行改进。

## 1 研究方法

### 1.1 跳层连接结构

#### 1) 注意力机制 CBAM

为了能够让模型同时关注图像的目标特征信息和全局特征信息,并解决信息过载的问题,本文在跳层连接中加入了 CBAM,具体结构如图 1 所示。该模型由通道注意力机制(channel attention module)和空间注意力机制(spatial attention module)组成。CBAM 的输入特征(Input)是编码器卷积(Conv2d)层的输出特征,通道注意力机制用来学习重要的图像特征信息,空间注意力机制用来学习全局上下文信息。最后将图像的全局信息和重要特征信息进行融合,得到输出特征向量(Output)。

注意力通道如图 1a 所示,首先将输入特征分别输入到最大池化层(Maxpool)和平均池化层中(Avgpool)中,得到两个特征图  $F_{AVG}$  和  $F_{Max}$ 。然后使用两层多层感知器 MLP 结构来实现神经元共享。最后,通过对 MLP 的输出特征进行加法运算和并使用 Sigmoid 激活函数来获得  $M_c(F)$ 。

空间注意力机制如图 1c 所示。输入特征分别被输入最大池化层(Maxpool)和平均池化层(Avgpool),通过计算可以获得这两个输出特征;将输出特征在尺寸方向上叠加之后,使用卷积核进行维度处理并获得空间注意力矩阵,空间注意力机制的卷积核大小设置为  $1 \times 1$ ;最后,使用 Sigmoid 激活函数来获得模块的输出结果  $M_s(F)$ 。

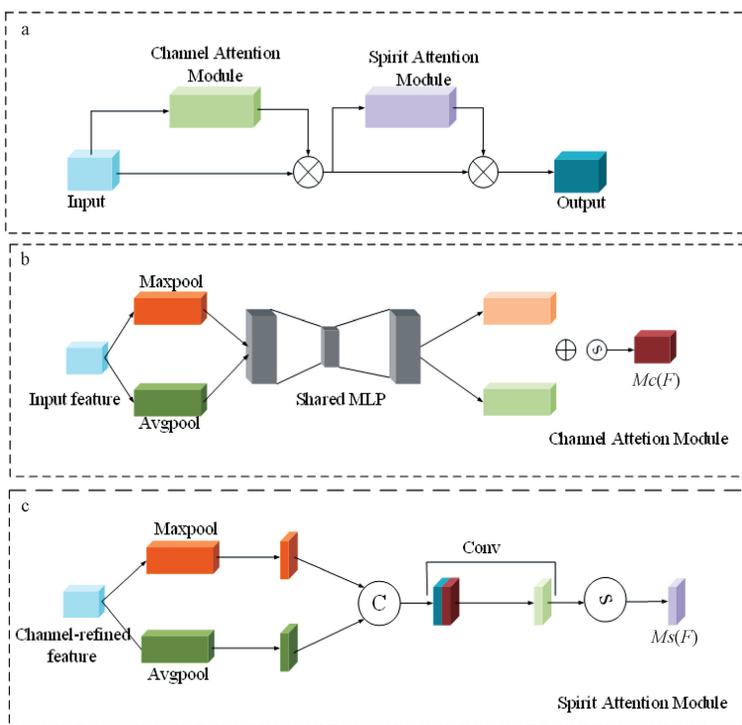


图 1 CBAM 模型结构

Fig. 1 CBAM model structure

2) 注意力特征融合模块 AFF

使用编码器下采样提取图像特征时,会丢失部分图像特征,导致重建图像和原始输入图像之间的差异增大,为了减少图像特征丢失并提高模型学习图像特征的能力,本文引入了注意力特征融合模块 (attentional feature fusion, AFF),目的是为了融合 CBAM 的输出特征向量和

解码器转置卷积 (ConvTranspose2d) 的输出特征向量,图 2 给出了 AFF 的模型结构。

图 2 中  $X_i$  是第  $i$  层转置卷积的输出特征向量,  $X_i$  和  $X_i$  维度相同的 CBAM 输出特征向量,为了保证 AFF 输入和输出图像特征的维度不变, AFF 中的卷积核大小全部采用  $1 \times 1$ , 步长全部设置为 1。

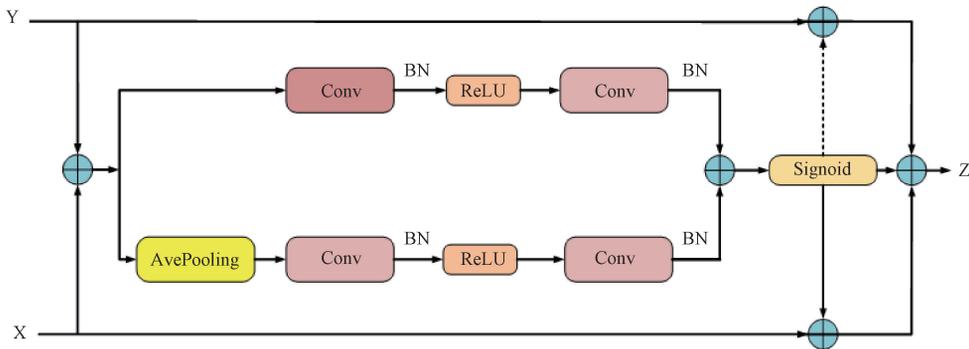


图 2 AFF 模型结构

Fig. 2 AFF model structure

1.2 联合上采样模块 JPU

当模型中加入 AFF 和 CBAM 之后,模型的结构变得更加复杂,导致模型的计算量增加。文献 [23] 中提出了 JPU (joint pyramid upsampling) 模型,并证明该方法可以降低模型的计算复杂度,并将运行速度提高

了 3 倍,因此为了能够提高模型的运行速度,本文在判别器中加入了 JPU 模块。图 3 给出了 JPU 的模型结构,在图 3 中 Conv2、Conv3 和 Conv4 分别连接判别器中卷积层的输出特征向量,具体的连接方法将会在 1.3 节中给出。

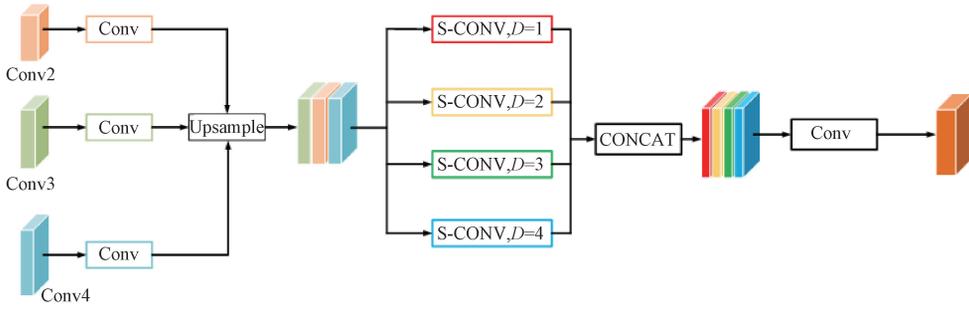


图 3 JPU 模型结构

Fig. 3 JPU model structure

### 1.3 整体网络框架

FAMGAN 的模型如图 4 所示。本文创新性的使用 CBAM 和 AFF 构建了新的跳连层,新的跳连层可以同时关注图像的目标特征信息和全局特征信息,不仅可以减少图像特征丢失,而且提高了模型学习图像特征的能力。判别器中加入了联合上采样模块 JPU,提高了模型的训练速度。图 3 中的 3 个卷积层 Conv2、Conv3 和 Conv4 分别与判别器的第 2 层、第 3 层和第 4 层卷积的输出特征相连。64、128、256、512 等数字表示输出特征层的维度。训练阶段使用编码器进行下采样获取更深层的图像特

征,编码器卷积的输出特征向量经过 CBAM 之后和转置卷积的输出特征向量一起输入 AFF 进行特征融合,利用融合之后的特征向量继续进行上采样,最终得到重建图像。

解码器每一层转置卷积输入特征向量如式(1)所示:

$$(Zq)_i = AFF(CBAM(Ze)_{i-1} + (Zt)_{i-1}) \quad (1)$$

式中:  $(Zq)_i$  表示第  $i$  层转置卷积的输入特征向量,  $(Zt)_{i-1}$  表示第  $i-1$  层转置卷积的输出特征向量,  $(Ze)_{i-1}$  表示和  $(Zt)_{i-1}$  维度相同的编码器卷积输出特征向量。

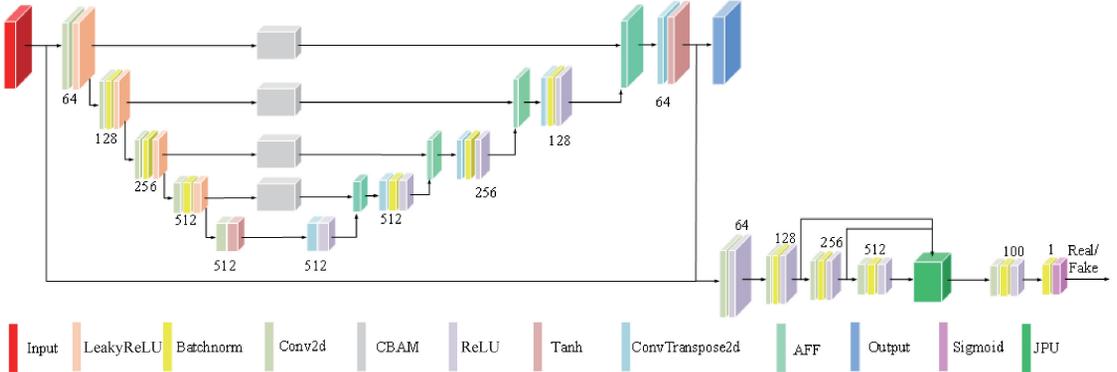


图 4 FAMGAN 网络模型

Fig. 4 FAMGAN network model

### 1.4 损失函数

为了能够让生成图像可以更接近原始输入图像, FAMGAN 采用了 3 个损失函数,包括重建损失,对抗损失和潜在损失,具体公式如下:

1) 重建损失:重建损失是原始输入图像和重建图像的差距,具体如式(2)所示。

$$L_{con} = E_{x \sim p_x} \| x - G(x) \|_1 \quad (2)$$

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。

2) 对抗损失:使用对抗损失目的是让生成器充分学习正常图像特征,迫使生成器重建与输入图像更加相似

的图像,对抗损失函数如式(3)所示。

$$L_{adv} = E_{x \sim p_x} \| D(x) - D(G(x)) \|_2 \quad (3)$$

3) 潜在损失:使用潜在损失函数目的是减少无关的信息干扰解码器重建图像,潜在损失函数如式(4)所示。

$$L_{lat} = E_{x \sim p_x} \| f(x) - f(G(x)) \|_2 \quad (4)$$

式中:  $f(\cdot)$  表示判别器最后一层卷积输出的潜在向量。

生成器总的损失函数如式(5)所示:

$$L_{total\_loss} = \omega_1 L_{con} + \omega_2 L_{adv} + \omega_3 L_{lat} \quad (5)$$

式中:  $\omega_1$ 、 $\omega_2$ 、 $\omega_3$  是 3 个损失函数对应的权重参数。

## 2 实验

### 2.1 自制数据集

本文使用的数据集是实验室自制的轮胎 X 光图像数据集, 每张图像的大小为  $512 \times 512$ , 缺陷包括气泡、胎稀、帘线弯曲和杂质 4 种类型。为了模拟实际工业生产过程中正常图像多, 异常图像少的特点, 在训练阶段本文使用 8 000 张无缺陷图像进行训练。模型训练完成之后, 使用测试集检测模型的训练效果, 测试集包括单类缺陷图像测试集和混合缺陷图像测试集。每一类数据集中缺陷图

像的数量如表 1 所示, 测试集中正常图像和异常图像的数量相同, 由于训练集图像数量足够多, 因此不再进行图像增强。部分正常图像和缺陷图像如图 5 所示, 图 5 中的缺陷图像依次是气泡、胎稀、帘线弯曲、杂质。

表 1 实验数据

Table 1 Experimental data

缺陷类型	图像数量
气泡	501
胎稀	455
杂质	513
帘线弯曲	442
混合缺陷	1 911

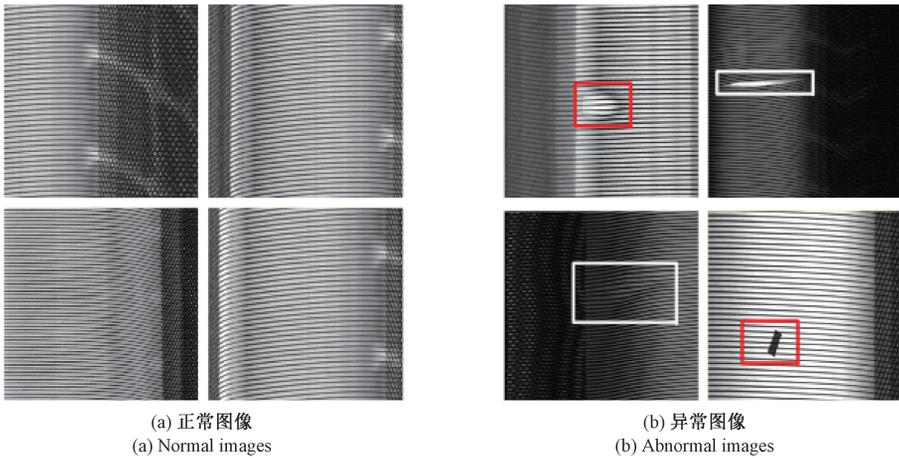


图 5 轮胎 X 光图像

Fig. 5 Tire X-ray image

### 2.2 实验配置

本次实验的开发平台配置为 Python3.9 和 Pytorch1.7.0, 深度卷积神经网络开发工具。计算机的硬件配置为 Inter i7 CPU、Nvidia GTX1080GPU、16 G 内存。编程环境为 Ubn-tu16.04LTS 操作系统。

FAMGAN 模型优化器采用 Adam<sup>[24]</sup>, 在进行训练之前先将输入图像由  $512 \times 512$  调整为  $128 \times 128$ , 具体的模型参数如表 2 所示。

表 2 参数设置

Table 2 Parameter settings

参数类型	参数设置
衰减率 $b_1$	0.5
衰减率 $b_2$	0.99
batch_size	24
学习率	0.000 2
$\omega_1$	40
$\omega_2$	1
$\omega_3$	1

模型训练需要大量的时间, 为了能够使用较少的训练次数得到较好的测试结果, 本文将 epoch 的值先后设置为 50、100、150 和 200, 得到的 AUC 值如表 3 所示。

表 3 实验结果

Table 3 Experimental results

epoch	50	100	150	200
气泡	0.837	0.779	0.849	0.812
杂质	0.872	0.807	0.864	0.838
胎稀	0.870	0.807	0.864	0.838
帘线弯曲	0.815	0.922	0.820	0.723
混合	0.890	0.873	0.895	0.842
平均	0.857	0.852	0.858	0.811

从表 3 可以看出, 将 epoch 设置为 50、100 和 150, FAMGAN 模型的 AUC 平均值没有太大变化。epoch 为 200 时, 模型的 AUC 平均值不理想, 综合考虑本文将 epoch 的值设置为 50。

### 2.3 实验结果和模型评估

本文使用的模型评估标准是 AUC (area under curve), AUC 是 ROC 曲线下边的面积, ROC 曲线的纵坐标是判断样本为正常图像的概率  $P_0$ , 横坐标表示判断为样本异常图像的概率  $P_1$ , 则 AUC 的值代表  $P_0 > P_1$  的概率。AUC 的值越大则分类越准确。

FAMGAN 和原始模型 (去掉跳层连接之后的 Skip-GANomaly) 的生成器损失函数如图 6 所示。

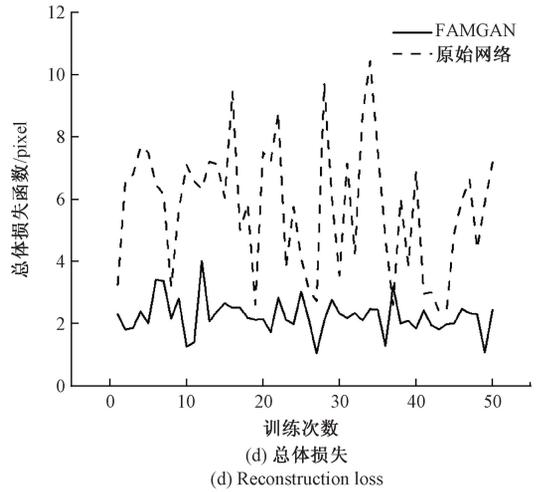
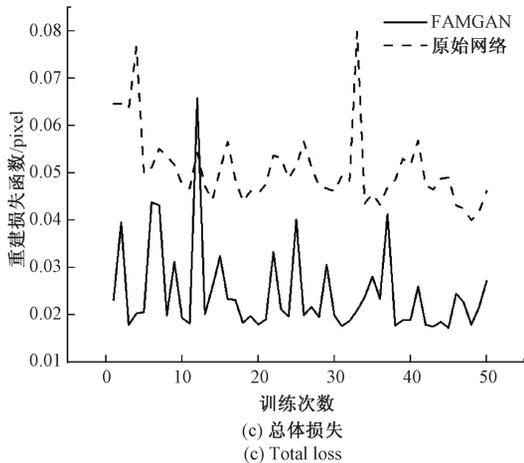
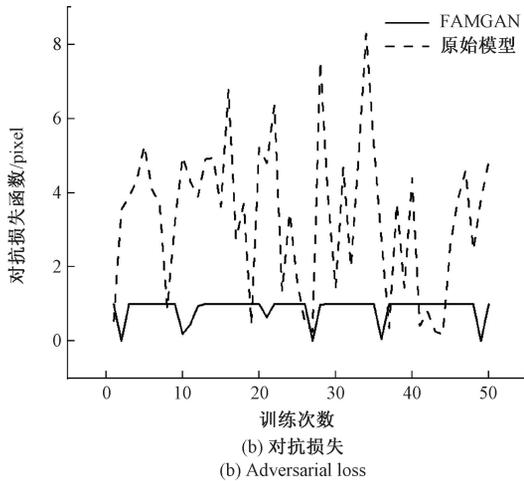
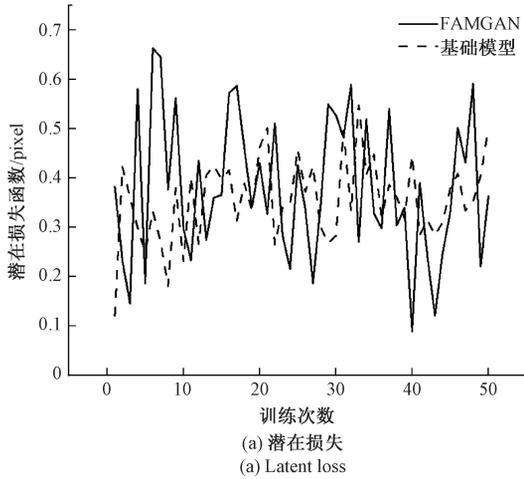


图 6 损失函数曲线

Fig. 6 Curve of loss function

由图 6 可以看出, FAMGAN 的重建损失函数、对抗损失函数和总体损失函数比原始模型都要小, FAMGAN 的重建损失基本上小于 0.04, 而原始模型的重建损失基本上大于 0.04, 证明与原始模型相比, FAMGAN 的重建图像更接近原始输入图像。

模型训练完成之后, 使用测试集检测模型的训练效果, 测试集中某一类缺陷的异常检测分数定义为:

$$A(x) = E_{x \sim p_x} \|x - G(x)\|_1 \quad (6)$$

为了评估模型的整体性能, 参考文献[25], 通过计算测试集中每个测试样本的异常分数值  $A(x)$ , 得到一组向量  $\mathbf{A} = \{a_i : A(x_i)\}$ , 并将  $\mathbf{A}$  中的每个值压缩到  $[0, 1]$  范围内, 每个测试样本异常值最后的更新结果如式 (7) 所示:

$$a_i = \frac{a_i - \min(\mathbf{A})}{\max(\mathbf{A}) - \min(\mathbf{A})} \quad (7)$$

### 2.4 对比实验

为了能够进一步证明 FAMGAN 模型和其他模型相比具有较好的优势。本文使用相同的数据集对 GANomaly、F-AnoGAN、AnoGAN、Skip-GANomaly 和 DEGAN 进行训练和测试。并将得到的实验结果进行对比, 具体结果如下。

#### 1) 测试时间对比

6 种网络模型的平均测试时间如表 4 所示, 从表 4 中可以看出, 在其他 5 种模型中 Skip-GANomaly 测试一张图像平均需要的时间最少, 本文构建的 FAMGAN 测试一张图像平均需要 67.04 ms, 比 Skip-GANomaly 少了 6.3 ms。

表 4 测试时间  
Table 4 Test time

网络模型	平均耗时/ms
F-AnoGAN	92.70
AnoGAN	173.73
GANomaly	101.79
Skip-GANomaly	73.34
FAMGAN	67.04
DEGAN	76.60

表 5 AUC 值  
Table 5 AUC values

模型类型	AnoGAN	F-AnoGAN	GANomaly	Skip-GANomaly	FAMGAN	DEGAN
气泡	0.738	0.704	0.763	0.530	0.837	0.447
胎稀	0.772	0.541	0.864	0.982	0.870	0.857
杂质	0.789	0.541	0.864	0.982	0.872	0.857
帘线弯曲	0.821	0.737	0.912	0.970	0.815	0.881
混合	0.787	0.672	0.858	0.793	0.890	0.716
平均	0.781	0.639	0.852	0.852	0.857	0.752

3) 图像重建和异常点定位对比

模型训练完成之后,由于模型只学习了正常的图像特征,测试阶段当缺陷图像通过生成器时,图像的缺陷特征会被丢弃,理论上可以得到与正常图像无任何差异的重建图像,通过计算原始图像和重建图像的差值可以得

2) 精度对比

5 种模型的 AUC 值如表 5 所示,F-AnoGAN 表现最差。FAMGAN 检测气泡缺陷得到的 AUC 为 0.837,与 GANomaly 相比提高了 7.4%。4 种缺陷类型混合检测得到的 AUC 值为 0.890,与 GANomaly 相比提高了 3.2%。FAMGAN 检测气泡和杂质缺陷类型得到的 AUC 值分别为 0.870 和 0.872,仅次于 Skip-GANomaly。FAMGAN 检测帘线弯曲时得到的 AUC 值为 0.815,与其他模型相比并不具有优势。

到缺陷图像对应的热力图,利用热力图可以准确找到图像缺陷的位置。图 7 展示了部分缺陷图像、重建图像以及热力图。从图 7 可以看出 FAMGAN 对于气泡的重建效果与其他网络模型相比有了明显的提高。

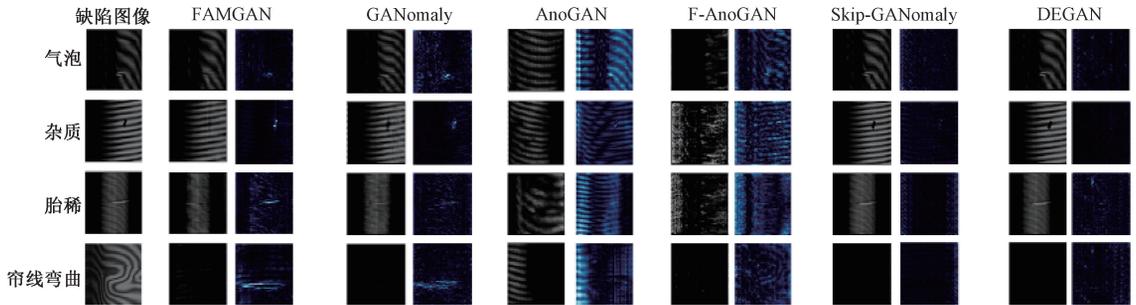


图 7 检测结果

Fig. 7 The detection results

2.5 消融实验

1) 第 1 类消融实验

本文构建的 FAMGAN 包括 CBAM、AFF 和 JPU,为了探究每个模块对实验结果的影响,本文进行了消融实验,最后得到的结果如表 6 所示。表 6 中的基础 GAN 是去

掉跳层连接之后的 Skip-GANomaly,通过表 6 可以看出单独使用 AFF 只对气泡的 AUC 值提高了 0.148。单独使用 CBAM 和单独使用 JPU 进行图像缺陷检测得到的结果,与基础 GAN 相比 AUC 的平均值分别提高了 0.142 和 0.085。

表 6 消融实验结果

Table 6 Results of ablation experiment

模型类型	基础 GAN	基础 GAN+AFF	基础 GAN+CBAM	基础 GAN+JPU	FAMGAN
帘线弯曲 AUC	0.671	0.653	0.849	0.765	0.815
气泡 AUC	0.623	0.771	0.810	0.835	0.837
胎稀 AUC	0.720	0.605	0.806	0.712	0.870
杂质 AUC	0.720	0.605	0.806	0.712	0.872
混合 AUC	0.701	0.736	0.874	0.834	0.890
平均 AUC	0.687	0.673	0.829	0.772	0.856

## 2) 第 2 类消融实验

为了验证不同模块联合使用对异常检测性能的影响,本文进行了第 2 类消融实验。第 2 类实验是将 3 个模块随机组合,实验的类型分别包括基础 GAN+AFF+JPU、基础 GAN+CBAM+JPU 和基础 GAN+AFF+CBAM。最后得到的结果如表 7 所示。基础 GAN+CBAM+AFF 的 AUC 平均值比基础 GAN+AFF+JPU 和基础 GAN+CBAM+JPU 分别提高了 0.054 和 0.007,尤其是检测气泡缺陷得到的 AUC 值比其他方法有了明显的提高。基础 GAN+AFF+CBAM 和 FAMGAN (基础 GAN+AFF+CBAM+JPU) 相比,AUC 的值并没有多大改变,但 FAMGAN 的平均耗时仅为 67.04 ms,基础 GAN+AFF+CBAM 的平均耗时 149.05 ms。证明将 CBAM、JPU 和 AFF 相结合可以更好的检测背景和缺陷较为相似的缺陷图像,同时可以提高模型的检测速度。

表 7 不同模块相结合得到的结果

Table 7 Results obtained by combining different modules

模型类型	基础 GAN +AFF+JPU	基础 GAN +CBAM+JPU	基础 GAN +AFF+CBAM	FAMGAN
帘线弯曲 AUC	0.874	0.921	0.815	0.815
气泡 AUC	0.606	0.628	0.840	0.837
胎稀 AUC	0.873	0.945	0.867	0.870
杂质 AUC	0.873	0.945	0.867	0.872
混合 AUC	0.773	0.798	0.882	0.890
平均 AUC	0.800	0.847	0.854	0.856
平均耗时 ms	106.63	75.10	149.05	67.04

## 3 结 论

本文使用去掉跳连层的 Skip-GANomaly 为基本框架,构建了融合注意力机制的生成对抗网络 FAMGAN, FAMGAN 中加入了 CBAM、AFF 和 JPU,使用 AFF 和 CBAM 构建了新的跳层连接,增强了模型学习图像特征和关注目标特征的能力,并在判别器中加入了 JPU 用来提高模型的检测速度。模型构建完成后,使用实验室的自制数据集进行训练和测试,FAMGAN 气泡图像的 AUC 值为 0.837,比 GANomaly 提高了 7.4%。FAMGAN 的 AUC 平均值与其他模型相比也有所提高。但帘线弯曲 FAMGAN 的检测结果依然不理想,下一步将继续探索如何提高这一类缺陷图像的检测准确率。

## 参考文献

[ 1 ] PSUJ G. Multi-sensor data integration using deep learning for characterization of defects in steel elements [ J ]. Sensors, 2018, 18(1): 292.

[ 2 ] GOPALAKRISHNAN K, KHAITAN S K, CHOUDHARY A, et al. Deep convolutional neural networks with

transfer learning for computer vision-based data driven pavement distress detection [ J ]. Construction and Building Materials, 2017, 157: 322-330.

- [ 3 ] LIN H, LI B, WANG X, et al. Automated defect inspection of LED chip using deep convolutional neural network [ J ]. Journal of Intelligent Manufacturing, 2019, 30: 2525-2534.
- [ 4 ] TABERNIK D, ŠELA S, SKVARČ J, et al. Segmentation based deep-learning approach for surface-defect detection [ J ]. Journal of Intelligent Manufacturing, 2020, 31(3): 759-776.
- [ 5 ] MIRZA M, OSINDERO S. Conditional generative adversarial nets [ J ]. arXiv preprint arXiv: 1411.1784, 2014.
- [ 6 ] 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. Cham: Springer International Publishing, 2017: 146-157.
- [ 7 ] 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.
- [ 8 ] AKCAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Ganomaly: Semi-supervised anomaly detection via adversarial training [ C ]. Computer Vision-ACCV 2018: 14th Asian Conference on Co-mputer Vision, Perth, Australia, December 2-6, 2018, Revised Selected Papers, Part III 14. Springer International Publishing, 2019: 622-637.
- [ 9 ] TANG T W, KUO W H, LAN J H, et al. Anomaly detection neural network with dual auto-encoders GAN and its industrial inspection applications [ J ]. Sensors, 2020, 20(12): 3336.
- [ 10 ] WANG Y, ZHANG Y, ZHENG L, et al. Unsupervised learning with generative adversarial network for automatic tire defect detection from X-ray images [ J ]. Sensors, 2021, 21(20): 67-73.
- [ 11 ] 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.
- [ 12 ] NIU S, LI B, WANG X, et al. Defect image sample generation with GAN for improving defect recognition [ J ]. IEEE Transactions on Automation Science and Engineering, 2020, 17(3): 1611-1622.

- [13] ALVARENGA T A, CARVALHO A L, HONORIO L M, et al. Detection and classification system for rail surface defects based on Eddy current [ J ]. *Sensors*, 2021, 21(23): 7937.
- [14] SULTANI W, CHEN C, SHAH M. Real-world anomaly detection in surveillance videos [ C ]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 6479-6488.
- [15] CHEN M, LI C, LI K, et al. Double encoder conditional GAN for facial expression synthesis [ C ]. 2018 37th Chinese Control Conference (CCC). *IEEE*, 2018: 9286-9291.
- [16] MEI S, YANG H, YIN Z. An unsupervised- learning-based approach for automated defect inspection on textured surfaces [ J ]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 67 ( 6 ): 1266-1277.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [ C ]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [ C ]. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5 - 9, 2015, Proceedings, Part III 18*. Springer International Publishing, 2015: 234-241.
- [19] AKÇAY S, ATAPOUR-ABARGHOUEI 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.
- [20] YU J, LIU J. Two-dimensional principal component analysis-based convolutional autoencoder for wafer map defect detection [ J ]. *IEEE Transactions on Industrial Electronics*, 2020, 68(9): 8789-8797.
- [21] DAI Y, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion [ C ]. *Proceedings of the IEEE-E/CVF Winter Conference on Applications of Computer Vision*, 2021: 3560-3569.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional

block attention module [ C ]. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.

- [23] WU H, ZHANG J, HUANG K, et al. Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation [ J ]. *arXiv preprint ar-Xiv: 1903.11816*, 2019.
- [24] CHEN L C, YANG Y, WANG J, et al. Attention to scale: Scale-aware semantic image segmentation [ C ]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 3640-3649.
- [25] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [ C ]. *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 618-626.

### 作者简介



刘韵婷, 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** received a B. Sc. 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.



刘鑫 (通信作者), 2020 年于长春建筑学院获学士学位, 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: 1529466456@qq.com

**Liu Xin** (Correspondence author) received a B. Sc. degree from Changchun University of Architecture and Civil Engineering in 2020. He is now a M. Sc. candidate of Shenyang Ligong University. His main research interest includes anomaly detection.



高宇, 2018 年于东北电力大学获学士学位, 现为沈阳理工大学硕士研究生, 主要研究方向为三维重建。

E-mail: 2930899746@qq.com

**Gao Yu** received a B. Sc. degree from Northeast Electric Power University in 2018. She is now a M. Sc. candidate of Shenyang Ligong University. Her main research interest includes iterative reconstruction.