

DOI: 10.13382/j.jemi.B2306805

# 家用燃气表外观缺陷的改进 ViT+FastFlow 检测方法研究\*

高泽铭<sup>1</sup> 刘桂雄<sup>1</sup> 陈国宇<sup>2</sup>

(1. 华南理工大学机械与汽车工程学院 广州 510640; 2. 广州能源检测研究院 广州 511447)

**摘要:**外观质量是家用燃气表(DGM)国家强制检定项目之一,针对DGM外观质量检定中匮乏缺陷样本使基于有监督学习检测方法难以泛化到实际应用场景问题,本文研究DGM外观缺陷无监督检测方法,引入Vision Transformer(ViT)改进版EfficientFormerV2-1提取正常样本特征,融合底层和高层特征图,并通过二维标准化流FastFlow将正常特征图映射到标准高斯分布,外观缺陷因离散落在分布以外使异常得分相比正常样本更高,通过设置自适应阈值识别并定位DGM外观缺陷。实验采集DGM正常样本、真实缺陷样本、合成缺陷样本作为数据集并优化检测模型参数,优化后检测模型在图像级别指标AUROC达99.77%,在像素级别指标AUPRO达96.3%,每秒可检测4张以上DGM图像,表明本文方法能准确高效识别与定位DGM外观缺陷。

**关键词:**外观缺陷;无监督学习;Vision Transformer;标准化流

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 460

## Research on the improved ViT+FastFlow detection method for appearance defects of domestic gas meters

Gao Zeming<sup>1</sup> Liu Guixiong<sup>1</sup> Chen Guoyu<sup>2</sup>

(1. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China;

2. Guangzhou Institute of Energy Testing, Guangzhou 511447, China)

**Abstract:** Appearance quality is one of the national mandatory verification for domestic gas meters (DGM). In view of the lack of defect samples in the appearance quality verification of DGM, which makes the detection method based on supervised learning difficult to generalize to the actual application scenario. This paper studies the unsupervised detection method of DGM appearance defects. EfficientFormerV2-1, the improved Vision Transformer (ViT), is introduced to extract normal sample features, fuse the bottom and high-level feature maps, and map the normal features to the standard Gaussian distribution using two-dimensional normalizing flow called FastFlow. The appearance defects are scattered outside the distribution so that the abnormal score is higher than the normal sample. By setting an adaptive threshold, the DGM appearance defects are identified and located. The experiment collects DGM normal samples, real defect samples, synthetic defect samples as data sets and optimizes the detection model parameters. The optimized detection model achieves 99.77% AUROC at image level indicators, 96.3% AUPRO at pixel level indicators, and can detect more than 4 DGM images per second, indicating that the method in this paper can accurately and efficiently identify and locate DGM appearance defects.

**Keywords:** appearance defects; unsupervised learning; Vision Transformer; normalizing flow

## 0 引言

DGM是重要民用计量仪表,国家标准GB/T 6968-2019、JJG 577-2012等对DGM外观质量做出明确要求,

但是外观缺陷种类繁多、缺陷形状、大小以及位置不固定、人工检定工作量极大。目前外观缺陷自动检测方法主要为传统机器视觉以及基于深度学习的检测方法。传统机器视觉检测方法通过图像处理方法或结合传统机器学习方法识别外观缺陷,如文献[1]提取缺陷纹理、灰

度、形态等特征,构建基于支持向量机的缺陷分类模型;文献[2]基于图像边缘特征将检测图像与合格图像配准,通过图像差分分割出缺陷;文献[3]通过灰度共生矩阵方法、频域变换、方向梯度直方图提取图像特征,并采用随机森林分类器对提取的缺陷特征进行分类。总之,传统机器视觉检测方法需人为设计外观缺陷匹配方法,泛化性较差。基于深度学习的检测方法由卷积神经网络(convolutional neural network, CNN)<sup>[4-5]</sup>或 ViT<sup>[6]</sup>等深度神经网络构成提取模型,能自动提取图像底层特征、高级语义信息,准确区分正常特征与异常特征,通用性强,其中有监督方法通过外观缺陷标签训练并优化检测模型特征提取网络、分类器,学习各类缺陷差异、同种缺陷相似之处,实现外观缺陷分类。如文献[7]研究多尺度特征图的目标检测改进方法,采用空洞卷积、层级特征融合算法准确快速地检测缺陷;文献[8]对特征提取主干网络融合多个残差块,并经 K-means 聚类算法,提高缺陷定位精度;文献[9]提出多支路卷积神经网络,并引入注意力模块突出图像的重要特征,通过级联融合各支路特征,提高缺陷检测准确率。有监督学习方法直接受学习样本量影响,工业场景下异常样本匮乏,限制其鲁棒性能,而无监督检测方法在训练过程中仅使用正常样本,在测试中可识别和定位外观缺陷,如文献[10-11]采用卷积自编码器重构无缺陷图像,并用重构无缺陷图像减去输入图像,获得包含缺陷信息的残差图;文献[12]训练图像修复网络模型恢复人为设置待检图像缺失区域内容,将重构图像前后明显差异点判别为缺陷;文献[13]提取最具代表性的正常样本特征点,构建特征向量记忆池来表征待检图像,使用最近邻方法计算待检图像表征后的异常分数图,设置阈值筛选缺陷区域;文献[14]使用标准化流建立正常特征空间高斯分布,分离待检图像的缺陷特征。

本文研究 DGM 外观缺陷无监督检测方法,引入 ViT 改进版 EfficientFormerV2-l<sup>[15]</sup>提取正常样本特征,并通过二维标准化流 FastFlow<sup>[16]</sup>映射到标准高斯分布,外观缺陷因离散落在分布以外使异常得分相比正常样本更高,通过设置自适应阈值识别并定位 DGM 外观缺陷,实现 DGM 外观缺陷检测的高效化。

## 1 基于改进 ViT+FastFlow 的 DGM 外观缺陷检测整体方案设计

DGM 外观缺陷种类繁多,除涂层磨损、结构断裂、表层划伤、结构破裂、印刷不良等常见缺陷,存在大量形状、大小各异未知多样外观缺陷。针对异常样本匮乏问题,本文采用基于改进 ViT+FastFlow 的 DGM 外观缺陷检测方法,利用正常样本表征其通用特性,并显性划分出正常

样本、异常样本类间差异。

图 1 为 DGM 外观缺陷检测模型结构图,主要包括特征提取模块(EfficientFormerV2)、概率分布变换模块(FastFlow)、自适应阈值分割模块等,其中, EfficientFormerV2 经大数据集训练后能有效获取 DGM 正常样本具有区分度的特征,其在前馈传递时能表征正常图像纹理、颜色、形状、位置等底层特征及各种高级语义信息,输出 DGM 正常样本不同尺度特征图。FastFlow 将图像特征划分为符合高斯分布的、易于处理的正常、异常区间,具体用可学习概率密度估计模型  $P, P = \{P_\theta; \theta \in \Theta\}$  在训练过程中学习参数空间  $\Theta(\theta)$  中参数  $\theta$ ,将图像特征  $x \in P_X(x)$  通过系列变换可逆映射到高斯分布空间  $z \in P_Z(z)$ <sup>[17-18]</sup>,其中  $\det(\cdot)$  为行列式:

$$P_X(x) = P_Z(z) \left| \det\left(\frac{\partial z}{\partial x}\right) \right| \quad (1)$$

对数似然函数  $\log P_X(x)$  表示  $P_X(x)$  与  $P_Z(z)$  分布相似程度<sup>[17-18]</sup>,训练过程中  $\log P_X(x)$  达到最大时,获得  $P$  的最优参数空间  $\Theta(\theta)$ :

$$z = f_\theta(x) \quad (2)$$

$$\log P_X(x) = \log P_Z(f_\theta(x)) + \log \left| \det\left(\frac{\partial f_\theta(x)}{\partial x}\right) \right| \quad (3)$$

外观缺陷不在正常特征分布空间,离散落在  $P_Z(z)$  分布以外,  $\log P_X(x)$  更低,其值可用于计算图像异常分数  $S$ ,则归一化对数似然值  $\bar{P}_X(x)$  为:

$$\bar{P}_X(x) = \exp(\log P_X(x) - \max[\log P_X(x)]) \quad (4)$$

$$S = \max[\bar{P}_X(x)] - \bar{P}_X(x) \quad (5)$$

设定阈值  $\tau$ ,若  $S > \tau$  则判定图像异常,避免人为设置干扰检测效果。

自适应阈值分割模块以检测性能指标为目标函数,优化过程中得出分割异常区域和正常区域的最佳阈值。推理阶段测试图像输入到已训练的 DGM 外观缺陷检测模型,当存在异常点时输出外观缺陷标签、外观缺陷位置及外观缺陷掩膜。

### 1.1 EfficientFormerV2

基于图像识别共通性,经大数据集(如 ImageNet)训练后的 CNN 或 ViT 深度神经网络能有效地获取 DGM 正常样本特有表征,是后续建立具有区分性分布、准确判别异常点的基础。

由于 CNN 能捕获位置等局部信息,但忽略图像全局与局部关系<sup>[19]</sup>;ViT 检测准确性更高,但缺乏归纳偏置能力,致模型结构复杂、运算低效<sup>[20]</sup>。为此,本文提出采用 EfficientFormerV2-l 轻量化混合架构(CNN+ViT)<sup>[12]</sup>提取图像特征,图 2 为 DGM 特征提取模块 EfficientFormerV2-l 模型结构图。

EfficientFormerV2-l 包括 Stem 模块、Local 模块、

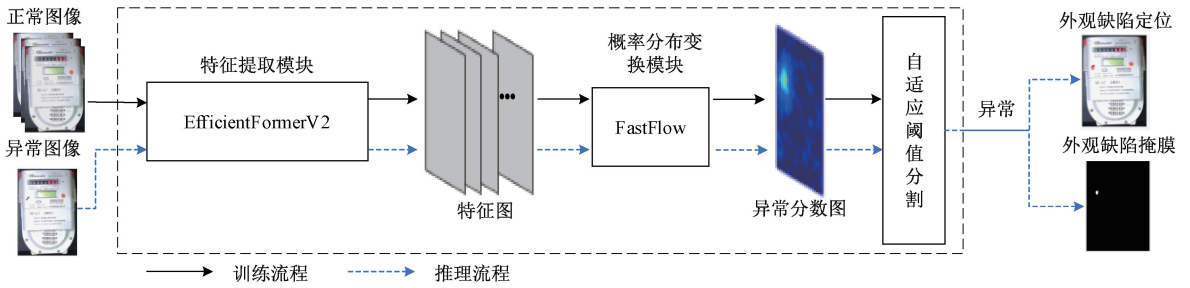


图 1 DGM 外观缺陷检测模型结构

Fig. 1 Structure of DGM appearance defect detection model

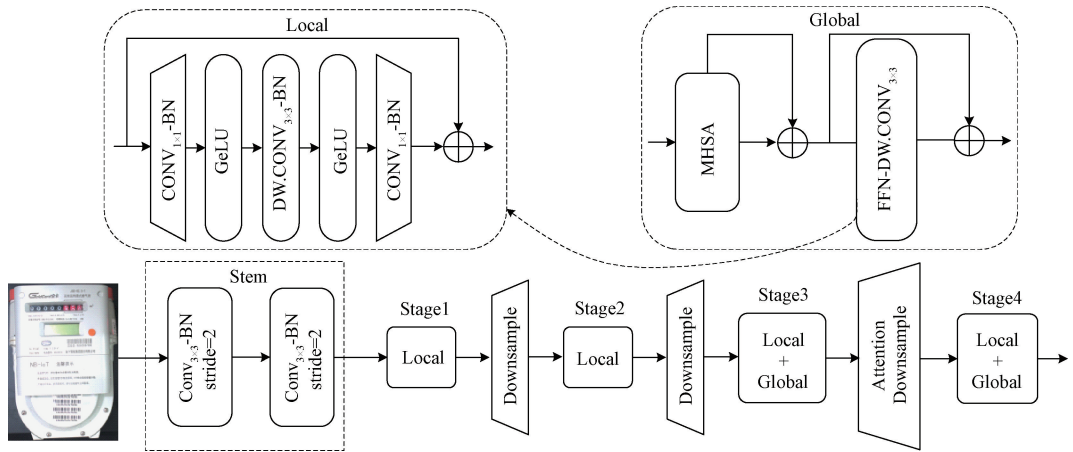


图 2 DGM 特征提取模块 EfficientFormerV2-1 模型结构

Fig. 2 Structure of DGM feature extraction module (EfficientFormerV2-1)

Downsample 模块、Global 模块、Attention Downsample 模块等,图 3 为 Global 模块、Attention Downsample 的 MHSA 结构图。

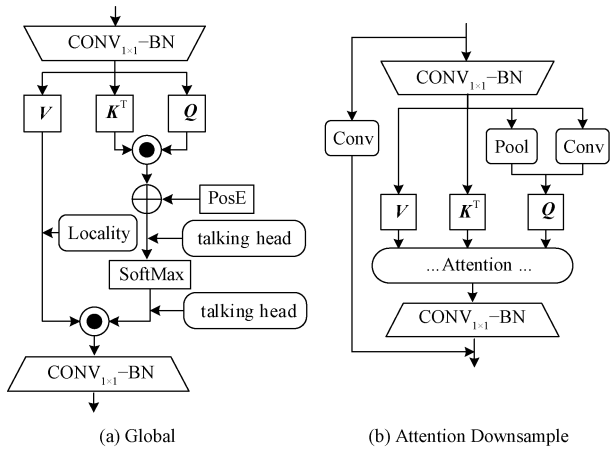


图 3 关键模块 MHSA 结构

Fig. 3 Key module MHSA structure

其中,Stem 模块由两个  $3 \times 3$  跨步为 2 卷积层组成,对输入图像下采样 4 倍,减少冗余信息;Local 模块采用

前馈神经网络 (feedforward network, FNN) 结构,中间插入深度可分离卷积 (depthwise convolution, DW. Conv) 提高局部特征提取能力;Downsample 模块采用跨步为 2 的卷积层,对输入特征图下采样;Global 模块通过添加  $3 \times 3$  卷积层将局部信息融合到多头自注意力结构 (multi-headed self-attention, MHSA) 值矩阵  $V$  中,在 softmax 操作前后添加全连接层,实现自注意力头部之间通信,提升全局建模能力;Attention Downsample 模块通过自注意力机制,实现全局上下文感知下采样,采用池化层 (pool)、跨步卷积层 (Conv) 局部下采样特征图,并将结果组合投影到查询矩阵  $Q$ ,使输出分辨率减半。同时,自注意力下采样模块、跨步卷积层 (Conv) 残差连接,融合局部与全局信息。

图 4 为 EfficientFormerV2-1 提取特征图可视化热图,Stage1、Stage2 低层网络能识别家用燃气表轮廓、形状、位置等底层信息,Stage3、Stage4 输出特征图具有高级语义信息,融合多个不同层特征图建立特征空间分布。

### 1.2 FastFlow

图 1 中 FastFlow 为二维标准化流 (normalizing flow, NF) 模型,它由如图 5 所示的多个叠加可逆转换模块组

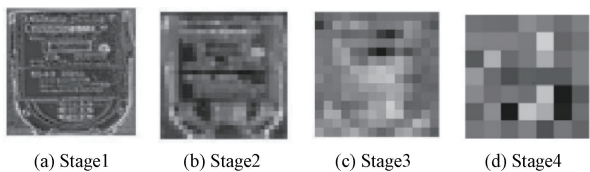


图 4 EfficientFormerV2-1 提取特征图可视化热图

Fig. 4 Visualize heat maps of extract feature maps (efficientFormerV2-1)

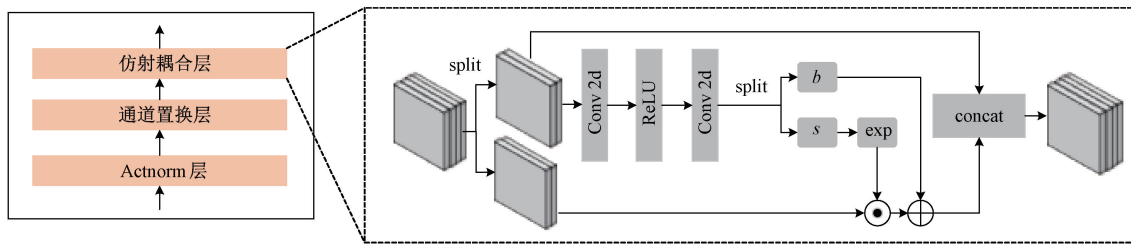


图 5 FastFlow 可逆转换模块结构

Fig. 5 Structure of reversible conversion module (FastFlow)

1) 输入特征图沿通道维度拆分为两部分,通过二维卷积层分离出仿射变换参数  $s, t$ :

$$\begin{cases} x_a, x_b = split(x) \\ (\log s, b) = CNN(y_b) \\ s = \exp(\log s) \end{cases} \quad (7)$$

2) 特征图一半做仿射变换,另一半保留原值,拼接两部分形成新分布的特征图:

$$\begin{cases} y_a = s \odot x_a + b \\ y_b = x_b \\ y = concat(y_a, y_b) \end{cases} \quad (8)$$

为筛选出正常样本最优高斯分布表征模型,采用梯度下降法训练 FastFlow 模型,使损失函数  $L_C$  达到最小:

$$L_C(\theta) = -\log P_X(x) = \frac{1}{N} \sum_{i=1}^N \left[ \frac{\|z_i\|_2^2}{2} - \log |\det J_i| \right] \quad (9)$$

其中,  $z_i$  为 FastFlow 输出,  $J_i$  为 FastFlow 输出与输入之间的雅克比矩阵,  $N$  为正常样本数。

图 6 为 FastFlow 正常特征和异常特征异常分数分布, FastFlow 能有效区分正常特征、异常特征,异常特征的异常分数更高,使用阈值  $\tau$  可判别出外观缺陷。

### 1.3 自适应阈值分割模块

为定位外观缺陷区域,将图像特征异常分数图插值上采样,复原到原图像大小,由每个像素异常得分分割出 DGM 外观缺陷区域。异常特征投影到正常样本高斯分布空间异常分数更高,利用阈值  $\tau$  能有效分割出家用燃气表外观缺陷。将正常样本的最大异常分数作为分割阈

成,将正常特征  $x \in P_X(x)$  映射到标准高斯分布空间  $P_Z(z)$  [12]:

$$P_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad (6)$$

Actnorm 层对每个通道进行仿射转换,使输入数据满足均值为 0、标准方差为 1 分布;通道置换层变换通道顺序,使每次迭代混合信息流;仿射耦合层采用二维卷积层保留流模型中空间信息,结合仿射转换学习正常特征潜在分布,具体转换如下:

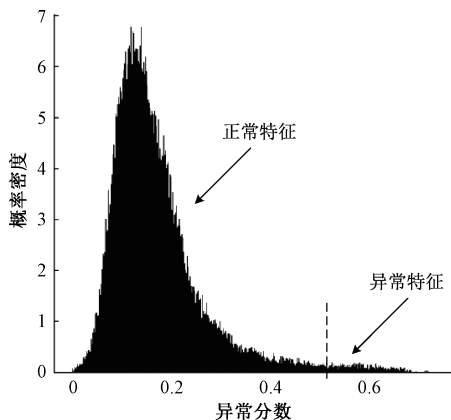


图 6 FastFlow 正常特征和异常特征异常分数分布  
Fig. 6 Normal feature and abnormal feature anomaly fraction distribution of FastFlow




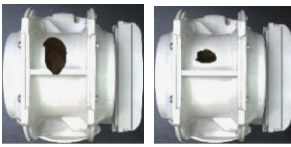
值,将会产生大量漏检,可借助异常样本自适应确定最佳阈值。

为扩充异常样本数据量,裁剪异常样本缺陷,可应用几何变换、色彩变换、高斯模糊等数据增强[21]生成多样缺陷,随机剪贴到正常图像不同位置合成异常图像,模拟异常图像特征分布,表 1 为多视图 DGM 部分外观缺陷合成图像。

正常图像、真实异常图像和合成异常图像集成一个测试数据集,输入到已训练好的 DGM 外观缺陷检测模型,在不同阈值  $\tau$  下得到外观缺陷检测  $F1score$  曲线,选择其中最佳阈值  $\bar{\tau}$  分割 DGM 外观缺陷:

表 1 多视图 DGM 部分外观缺陷合成图像

Table 1 Part of Multi-view DGM appearance defect composite images

视图	正常样本	缺陷类型	合成异常图像
正视		塑料破裂	
仰视		金属破裂	

$$\begin{cases} precision = \frac{TP}{TP + FP} \\ recall = \frac{TP}{TP + FN} \\ \max(F1score) = \frac{2 \times precision \times recall}{precision + recall} (\bar{\tau}) \end{cases} \quad (10)$$

其中,  $TP$  为正确预测外观缺陷,  $FP$  为错将正常类预测为外观缺陷,  $FN$  为错将外观缺陷预测为正常类,  $precision$  为精确率,  $recall$  为查全率。

## 2 实验与分析

家用燃气表正视方向能采集注塑件外观图像, 仰视方向能采集金属件外观图像, 以上视角采集的图像包含大多数家用燃气表外观缺陷种类, 采集家用燃气表正视图、仰视图各 200 张(其中每个视图各有 10 张为异常)作为数据集, 并使用改变光照、色调调换、仿射变换等数据增强方法扩大数据集。实验采用英伟达 GeForce RTX 3090 GPU 训练 DGM 外观缺陷检测模型、识别与定位 DGM 外观缺陷, 训练过程使用 Adam 优化器, 其学习率设置为  $10^{-3}$ , 权重衰减系数设为  $10^{-5}$ 。检测阶段采用家用燃气表外观缺陷评价指标评估不同检测模型性能, 并框选出其外观缺陷位置。

### 2.1 外观缺陷评价指标

1) AUROC (area under receiver operator characteristic curve) 分类指标为在不同阈值设置下查全率 (true positive rate, TPR) 与误检率 (false drop rate, FPR) 关系曲线下面积:

$$\begin{cases} TPR = recall = \frac{TP}{TP + FN} \\ FPR = \frac{FP}{FP + TN} \end{cases} \quad (11)$$

式中:  $TN$  为正确判断为正常样本。

2) AUPRO (area under per region overlap) 分割指标为在不同阈值设置下 PRO 与 FPR 关系曲线下面积:

$$PRO = \frac{1}{M} \sum_{m=1}^M \frac{TP_n}{TP_n + FN_n}, FPR \in (0, 0.3) \quad (12)$$

式中:  $M$  为预测区域和真实缺陷区域按照连通域划分区域个数。

3) FPS ((frame per second)) 用来评估外观缺陷检测速度, 测试检测模型每秒预测图片数量。

### 2.2 实验结果与分析

检测模型性能主要受特征提取模块及 FastFlow 结构、特征融合程度影响, 因此针对特征提取模块类别、FastFlow 堆叠数、特征融合方式分别进行消融实验, 得出本文综合性能最优的检测模型。

#### 1) 检测模型结构优化验证

依据文献[16]选取 Wide-ResNet50-2 (WR50)、DeiT-base-distilled (DeiT), 与 EfficientFormerV2-1 (EFV2-1) 做对比实验。表 2 为不同特征提取模块对检测外观缺陷检测性能影响, 结果表明 EFV2-1 兼具 CNN、ViT 特点, 相比其他骨干图像级 AUROC 最高达 99.77%, 更能准确分类正常点、异常点; EFV2-1 在像素级 AUPRO 指标较 WR50 提高 0.96%, 较 DeiT 降低 0.88%, 虽没有最优细粒化, 但模型结构轻量化, 参数量比 WR50、DeiT 模分别减少 69.1%、75.1%; 本文外观缺陷检测效率最高, FPS 达 4.25。因此本文采用 EFV2-1 具有最佳综合外观缺陷检测性能。

表 2 不同特征提取模块对检测外观缺陷检测性能影响  
Table 2 Influence of different feature extraction modules on the performance of appearance defect detection

骨干	结构	图像级	像素级	模型参数	
		AUROC/%	AUPRO/%	量/M	FPS
WR50	CNN	99.33	95.46	91.9	4.04
DeiT	ViT	98.59	97.18	114	3.4
EFV2-1	CNN+ViT	99.77	96.30	28.4	4.25

注: 方框中数据代表本列最优值, 下同

图 7 为 FastFlow 堆叠数与检测准确性关系实验结果, 验证过程中其他变量保持一致。由图 7 可知 FastFlow 堆叠数 > 0 的模型检测准确性远高于 FastFlow 堆叠数 = 0 的模型, 表明 FastFlow 有效构建正常样本分布空间并区分出外观缺陷, 提高检测模型准确性。当 FastFlow 堆叠数 = 8 时检测准确性最优, 图像级 AUROC 达 99.77%, 像素级 AUPRO 达 96.30%。

#### 2) 特征融合对检测模型影响验证

表 3 为不同层融合对外观缺陷检测性能影响, 实验过程中特征提取模块为 EFV2-1, 其他变量保持一致。



### 3 结 论

本文面向家用燃气表外观质量检定中缺陷样本匮乏问题,研究一种家用燃气表外观缺陷的改进 ViT+FastFlow 无监督检测方法,使用正常样本训练检测模型,仅需少量异常样本微调外观缺陷判定阈值,大大减少缺陷样本使用。将 EfficientFormerV2-l 作为正常图像特征提取网络,采用 FastFlow 将图像提取特征投射到标准高斯分布特征空间,外观缺陷因离散落在分布以外使异常得分相比正常样本更高,通过设置自适应阈值识别并定位 DGM 外观缺陷。检测模型在图像级别指标 AUROC 达 99.77%,在像素级别指标 AUPRO 达 96.3%,每秒可检测 4 张以上 DGM 图像,满足 DGM 外观缺陷准确性、实时性要求。本文虽然能识别并定位 DGM 外观缺陷,但是不能精细分类 DGM 各种未知多样外观缺陷,未来将针对该问题进行下一步工作。

### 参考文献

- [ 1 ] GAO M, WANG X, ZHU S, et al. Detection and segmentation of cement concrete pavement pothole based on image processing technology [ J ]. Mathematical Problems in Engineering, 2020, 2020: 1-13.
- [ 2 ] 谢晖,吴先锋,付山,等.基于 Halcon 的电表铭牌缺陷检测系统[J].现代制造工程,2021(2):119-126.
- XIE H, WU X F, FU SH, et al. Meter nameplate defect detection system based on Halcon [ J ]. Modern Manufacturing Engineering, 2021(2): 119-126.
- [ 3 ] CHAUDHARI C V. Steel surface defect detection using glcm, gabor wavelet, hog, and random forest classifier[J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(12): 263-273.
- [ 4 ] 彭伟康,陈爱军,吴东明,等.基于改进 Faster R-CNN 的水准泡缺陷检测方法[J].中国测试,2021,47(7):6-12.
- PENG W K, CHEN AI J, WU D M, et al. Defect detection method of level bubble based on improved Faster R-CNN[J]. China Measurement & Test, 2021, 47(7): 6-12.
- [ 5 ] 伍济钢,成远,邵俊,等.基于改进 YOLOv4 算法的 PCB 缺陷检测研究[J].仪器仪表学报,2021,42(10):171-178.
- WU J G, CHENG Y, SHAO J, et al. A defect detection method for PCB based on the improved YOLOv4 [ J ]. Chinese Journal of Scientific Instrument, 2021, 42(10): 171-178.
- [ 6 ] 高泽铭,刘桂雄,陈国宇,等.基于 ViT 燃气表外观零件识别与定位方法研究[J].电子测量技术,2023,

46(11):7-12.

GAO Z M, LIU G X, CHEN G Y, et al. Research on recognition and localization inspection of appearance for parts of gas meter based on ViT[J]. Journal of Electronic Measurement and Instrumentation, 2023,46(11): 7-12.

- [ 7 ] 李兰,奚舒舒,张才宝,等.基于改进 SSD 模型的工作表面缺陷识别算法[J].计算机工程与科学,2020,42(9):1608-1615.
- LI L, XI SH SH, ZHANG C B, et al. A surface defect recognition algorithm based on improved SSD model[J]. Computer Engineering & Science, 2020, 42(9): 1608-1615.
- [ 8 ] 吴伟浩,李青.基于改进 YOLO v3 的电连接器缺陷检测[J].传感技术学报,2020,33(2):299-307.
- WU W H, LI Q. Defect detection of electrical connector based on improved YOLO v3 [ J ]. Chinese Journal of Sensors and Actuators, 2020,33(2): 299-307.
- [ 9 ] 刘培勇,董洁,谢罗峰,等.基于多支路卷积神经网络的磁瓦表面缺陷检测算法[J].吉林大学学报(工学版),2023,53(5):1449-1457.
- LIU P Y, DONG J, XIE L F, et al. Surface defect detection algorithm of magnetic tiles based on multi-branch convolutional neural network[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(5): 1449-1457.
- [ 10 ] 罗月童,卞景帅,张蒙,等.基于卷积去噪自编码器的芯片表面弱缺陷检测方法[J].计算机科学,2020,47(2):118-125.
- LUO Y T, BIAN J SH, ZHANG M, et al. Detection method of chip surface weak defect based on convolution denoising auto-encoders [ J ]. Computer Science, 2020, 47(2): 118-125.
- [ 11 ] 余文勇,张阳,姚海明,等.基于轻量化重构网络的表面缺陷视觉检测[J].自动化学报,2020,48(9):2175-2186.
- YU W Y, ZHANG Y, YAO H M, et al. Visual inspection of surface defects based on lightweight reconstruction network [ J ]. Acta Automatica Sinica, 2020,48(9):2175-2186.
- [ 12 ] 胡广华,王宁,何文亮,等.基于图像修复的无监督表面缺陷检测方法[J].华南理工大学学报(自然科学版),2021,49(7):76-85.
- HU G H, WANG N, HE W L, et al. Unsupervised surface defect detection method based on image inpainting [ J ]. Journal of South China University of Technology (Natural Science Edition), 2021,49(7): 76-85.
- [ 13 ] 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.
- [14] GUDOVSKIY D, ISHIZAKA S, KOZUKA K. Cflow-ad: Real-time unsupervised anomaly detection with localization via conditional normalizing flows [ C ]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022; 98-107.
- [15] LI Y, HU J, WEN Y, et al. Rethinking vision transformers for mobilenet size and speed [ J ]. arXiv preprint arXiv: 2212. 08059, 2022.
- [16] YU J, ZHENG Y, WANG X, et al. FastFlow: Unsupervised anomaly detection and localization via 2D normalizing flows[J]. arXiv preprint arXiv:2111.07677, 2021.
- [17] DINH L, SOHL-DICKSTEIN J, BENGIO S. Density estimation using real nvp [ J ]. arXiv preprint arXiv: 1605.08803, 2016.
- [18] KINGMA D P, DHARIWAL P. Glow: Generative flow with invertible 1×1 convolutions[J]. Advances in Neural Information Processing Systems, 2018, 31.
- [19] RAGHU M, UNTERTHINER T, KORNBLITH S, et al. Do vision transformers see like convolutional neural networks? [ J ]. Advances in Neural Information Processing Systems, 2021, 34; 12116-12128.
- [20] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words; Transformers for image

recognition at scale [ J ]. arXiv preprint arXiv: 2010.11929, 2020.

- [21] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 1-48.

### 作者简介



**高泽铭**, 2018 年于华南理工大学获得学士学位, 现为华南理工大学在读研究生, 主要研究方向为现代传感技术与机器视觉。  
E-mail: megao\_zm@mail.scut.edu.cn

**Gao Zeming** received his B. Sc. degree from South China University of Technology in 2018. Now he is a M. Sc. candidate in South China University of Technology. His main research interests include modern sensing technology and machine vision.



**刘桂雄**(通信作者), 1990 年于重庆大学获得学士学位, 1995 年于重庆大学获得博士学位, 现为华南理工大学教授、博导, 主要研究方向为先进传感与仪器研究。

E-mail: megxliu@scut.edu.cn

**Liu Guixiong** (Corresponding author) received his B. Sc. degree from Chongqing University in 1990 and Ph. D. degree from Chongqing University in 1995, respectively. Now he is a professor and doctoral supervisor in South China University of Technology. His main research interests include advanced sensing and instrumentation research.