

同行专家业内评价意见书编号：20250854417

附件1

浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名：白玉虎

学号：22260371

申报工程师职称专业类别（领域）：电子信息

浙江工程师学院（浙江大学工程师学院）制

2025年03月21日

填表说明

一、本报告中相关的技术或数据如涉及知识产权保护、军工项目保密等内容，请作脱密处理。

二、请用宋体小四字号撰写本报告，可另行附页或增加页数，A4纸双面打印。

三、表中所涉及的签名都必须用蓝、黑色墨水笔，亲笔签名或签字章，不可以打印代替。

四、同行专家业内评价意见书编号由工程师学院填写，编号规则为：年份4位+申报工程师职称专业类别(领域)4位+流水号3位，共11位。

一、个人申报

（一）基本情况【围绕《浙江工程师学院（浙江大学工程师学院）工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】

1. 对本专业基础理论知识和专业技术知识掌握情况(不少于200字)

在工业图像异常检测领域的研究工作中，我系统掌握了该方向的理论体系与技术框架。基础理论方面，我具备一定的机器学习和计算机视觉理论基础，了解监督学习、无监督学习范式，特别针对异常检测场景中的小样本、高维度特性。阅读了大量关于异常检测的文献，对该领域内的方法有了一定的系统性地认识，熟悉常见的异常检测算法，如基于密度的方法、基于距离的方法、基于模型的方法等。

技术实践层面，在深度学习方法上，重点学习了自编码器（AE）、生成对抗网络（GAN）等无监督架构，能针对工业缺陷的局部特性改进模型注意力机制。具备一定的PyTorch框架下的模型开发能力，成功实现基于多尺度特征融合的卷积神经网络（CNN）和Vision Transformer（ViT）的混合架构，学会将理论的深度学习知识应用到实际中，了解多种深度学习的神经网络，

学习构建和训练不同类型的神经网络使得相关能力有所提升。通过实践提升了技能，包括编程能力和数据分析能力，通过使用python、pytorch进行代码编写使得编程能力得到一定程度的提升。

2. 工程实践的经历(不少于200字)

在面向实际工业场景的异常检测算法开发实践中，针对缺陷样本稀缺、标注成本高及异常形态多样性等核心挑战，我主导研发了一套基于频域特征增强的小样本异常检测框架。该项目的核心目标是在仅使用正常样本训练的条件下，实现对多品类工业产品中未知缺陷的精准检测与定位，同时突破传统方法在微小异常敏感度与跨产品泛化能力上的瓶颈。在多个工业检测场景的验证中，该方案在分别仅使用1, 2, 4张正常样本训练的条件下，取得平均AUROC 93.3%、95.7%、96.9%的分类精度和96.2%、97.2、97.5%的定位精度。

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例（不少于1000字）

我的研究课题为工业场景的异常检测，在实际工作中的面临的挑战和问题如下：在实际生产中获取大量带有缺陷的样本并对其进行准确标注是一项挑战。工业产品种类繁多，每种产品的缺陷类型、形状、大小和模式都可能不同，而我们在训练过程中未使用异常图像，如何在测试时准确检测并定位多种多样的异常。具体问题包括：1) 当前的小样本异常检测方法依旧无法高效地利用有限的样本；2) 它们可能无法检测和定位空间域中不明显的异常；3) 基于特征匹配的方法会受到预训练模型的影响（通常在ImageNet上进行预训练，而ImageNet

的数据与工业图像的特征分布存在偏差），以及这类方法需要额外的内存空间存储特征。具体的解决方法如下：首先，通过图像处理中简单的数据增强方式扩充数据集（随机旋转一定的角度即可取得不错的效果），以充分利用有限的样本。其次，一些微小的异常在频率域中会更容易被察觉。频域用于表示图像中各种模式与波动的频率和振幅。频域解耦的核心在于傅里叶变换及其相关理论。具体而言，傅里叶变换可将图像分解为一系列正弦分量，并通过其振幅和相位在频域中表征图像。因此可以将图像分解为不同频率的图像进行后续操作，高频图像通常记录细节纹理信息，而低频图像通常表示语义信息，因此可以通过图像处理中的频率分解将获取的图像分解为高频图像和低频图像，更加全面地进行信息分析。然后，使用一个特征适配器（深度学习中的全连接层）来缓解领域偏差，将正常特征和正常特征聚集到一起，同时将正常特征和异常特征进行区分，使其能够存在明显的边界。最后，由于提取的

适应性特征保留了丰富的空间信息，正常样本与异常样本在特征空间中的分布存在显著差异。直观上，可以将异常检测问题转换为特征空间中的分类问题，通过直接估计特征向量的异常性来实现异常检测那么可以采用一个判别器（由深度学习中全连接层和Vision Transformer构造的混合架构）在特征空间直接分辨正常特征和异常特征，而无需额外的存储空间储存特征。但是，仅使用正常图像进行训练容易造成过拟合现象，而且由于缺少正样本可能导致判别器和特征适配器无法进行优化，故将在图像层面和特征层面合成伪异常以监督整个网络进行训练，从视觉直观的角度出发，可以在图像层面构造异常，直接可以观察到异常的存在；而在另一方面，可以从特征层面构造异常，使得正常和异常特征分布存在明显的边界。尽管合成的异常与真实生产场景的异常存在不同，但是只需要其与正常的特征分布不同就可以帮助判别器识别异常，判别器在特征空间中学习异常特征和正常特征的联合表示。根据具体的情况设计了不同的模块，因此也需要具体的损失函数对不同的神经网络进行约束，使其能够达到预期的目标。通过设计三个损失函数对整个模型进行约束。通过上述一系列的操作，可以较好的解决实践中遇到的一些问题，并且能够较为准确的检测和定位异常。

(二) 取得的业绩(代表作)【限填3项, 须提交证明原件(包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等)供核实, 并提供复印件一份】

1. 公开成果代表作【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】

成果名称	成果类别 [含论文、授权专利(含发明专利申请)、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/授权或申请时间等	刊物名称/专利授权或申请号等	本人排名/总人数	备注
Dual-path Frequency Discriminators for few-shot anomaly detection	国际期刊	2024年08月15日	Knowledge-Based Systems	1/6	SCI期刊收录
PatchCLIP: Enhancing CLIP with Pair-images for Zero-shot Anomaly Detection	会议论文	2024年09月20日	Chinese Automation Congress	1/7	

2. 其他代表作【主持或参与的课题研究项目、科技成果应用转化推广、企业技术难题解决方案、自主研发设计的产品或样机、技术报告、设计图纸、软课题研究报告、可行性研究报告、规划设计方案、施工或调试报告、工程实验、技术培训教材、推动行业发展中发挥的作用及取得的经济社会效益等】

(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩： 84 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1.1 年(要求1年及以上) 考核成绩： 83 分
本人承诺	
<p>个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！</p> <p style="text-align: right;">申报人签名：白玉虎</p>	

浙江大学研究生院
攻读硕士学位研究生成绩单

学号: 22260371	姓名: 白玉虎	性别: 男	学院: 工程师学院	专业: 控制工程	学制: 2.5年						
毕业时最低应获: 25.0学分	已获得: 28.0学分			入学年月: 2022-09	毕业年月: 2025-03						
学位证书号: 1033532025602208			毕业证书号: 103351202502600196			授予学位: 电子信息硕士					
学习时间	课程名称	备注	学分	成绩	课程性质	学习时间	课程名称	备注	学分	成绩	课程性质
2022-2023学年秋季学期	研究生英语能力提升		1.0	免修	跨专业课	2022-2023学年秋冬学期	高阶工程认知实践		3.0	92	专业学位课
2022-2023学年秋季学期	创新设计方法		2.0	通过	专业选修课	2022-2023学年冬季学期	产业技术发展前沿		1.5	87	专业学位课
2022-2023学年秋季学期	智能无人系统及应用实践		2.0	80	专业选修课	2022-2023学年冬季学期	机器视觉及其应用		2.0	80	专业学位课
2022-2023学年秋季学期	研究生英语		2.0	免修	专业学位课	2022-2023学年冬季学期	智能控制技术		2.0	85	专业学位课
2022-2023学年秋季学期	工程数值分析		2.0	83	专业选修课	2022-2023学年春季学期	新时代中国特色社会主义思想理论与实践		2.0	80	专业学位课
2022-2023学年秋季学期	工程伦理		2.0	88	专业学位课	2022-2023学年春季学期	自然辩证法概论		1.0	91	公共学位课
2022-2023学年秋季学期	工程技术创新前沿		1.5	86	专业学位课	2022-2023学年春季学期	研究生论文写作指导		1.0	91	专业选修课
2022-2023学年秋季学期	研究生英语基础技能		1.0	免修	公共学位课		硕士生读书报告		2.0	通过	

说明: 1. 研究生课程按三种方法计分: 百分制, 两级制 (通过、不通过), 五级制 (优、良、中、及格、不及格)。
2. 备注中 "*" 表示重修课程。

学院成绩校核章:
成绩校核人: 张梦依
打印日期: 2025-03-31





Dual-path Frequency Discriminators for few-shot anomaly detection

Yuhu Bai ^{a b 1}✉, Jiangning Zhang ^{c d 1}✉, Zhaofeng Chen ^e, Yuhang Dong ^{a b}, Yunkang Cao ^f, Guanzhong Tian ^a ✉

Show more

Outline | Share Cite

<https://doi.org/10.1016/j.knosys.2024.112397>

[Get rights and content](#)

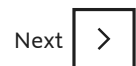
Full text access

Highlights

- Novel few-shot anomaly detection framework with four components.
- Constructing a dual-path discriminator from frequency perspective for learning joint representation is the significant components.
- Better performance than current state-of-the-art methods on benchmarks.

Abstract

Few-shot anomaly detection (FSAD) plays a crucial role in industrial manufacturing. However, existing FSAD methods encounter difficulties leveraging a limited number of normal samples, frequently failing to detect and locate inconspicuous anomalies in the spatial domain. We have further discovered that these subtle anomalies would be more noticeable in the frequency domain. In this paper, we propose a Dual-Path Frequency Discriminators (DFD) network from a frequency perspective to tackle these issues. The original spatial images are transformed into multi-frequency images, making them more conducive to the tailored discriminators in detecting anomalies. Additionally, the discriminators learn a joint representation with forms of pseudo-anomalies. Extensive experiments conducted on MVTec AD and VisA benchmarks demonstrate that our DFD surpasses current state-of-the-art methods. The code is available at <https://github.com/yuhbai/DFD> .



Keywords



经检索《Web of Science》、《Journal Citation Reports (JCR)》及《中国科学院文献情报中心期刊分区表》数据库,《Science Citation Index Expanded (SCI-EXPANDED)》收录论文及其期刊影响因子、分区情况如下。(检索时间:2024年12月26日)

第1条,共1条

标题: Dual-path Frequency Discriminators for few-shot anomaly detection

作者: Bai, YH(Bai, Yuhu);Zhang, JN(Zhang, Jiangning);Chen, ZF(Chen, Zhaofeng);Dong, YH(Dong, Yuhang);Cao, YK(Cao, Yunkang);Tian, GZ(Tian, Guanzhong);

来源出版物: KNOWLEDGE-BASED SYSTEMS 卷:302 文献号:112397 提前访问日期:AUG 2024

DOI:10.1016/j.knosys.2024.112397 出版年:OCT 25 2024

入藏号:WOS:001301703500001

文献类型:Article

地址:

[Bai, Yuhu; Dong, Yuhang; Tian, Guanzhong] Zhejiang Univ, Ningbo Innovat Ctr, Ningbo 315100, Peoples R China.

[Bai, Yuhu; Dong, Yuhang] Zhejiang Univ, Polytech Inst, Hangzhou 310015, Peoples R China.

[Zhang, Jiangning] Zhejiang Univ, Coll Control Sci & Engn, Hangzhou 310027, Peoples R China.

[Zhang, Jiangning] Tencent, YouTu Lab, Shanghai 200233, Peoples R China.

[Chen, Zhaofeng] China Tower Hangzhou Sci & Technol Innovat Ctr, Hangzhou 310020, Peoples R China.

[Cao, Yunkang] Huazhong Univ Sci & Technol, Sch Mech Sci & Engn, Wuhan 430000, Peoples R China.

通讯作者地址:

Tian, GZ (corresponding author), Zhejiang Univ, Ningbo Innovat Ctr, Ningbo 315100, Peoples R China.

电子邮件地址:yhbai@zju.edu.cn; 186368@zju.edu.cn; gztian@zju.edu.cn

IDS号:E2Y3V

ISSN:0950-7051

eISSN:1872-7409

期刊《KNOWLEDGE-BASED SYSTEMS》2023年的影响因子为7.2,五年影响因子为7.4。

期刊《KNOWLEDGE-BASED SYSTEMS》2023年的JCR分区情况为:

Edition	JCR® 类别	类别中的排序	JCR 分区
SCIE	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	27/197	Q1

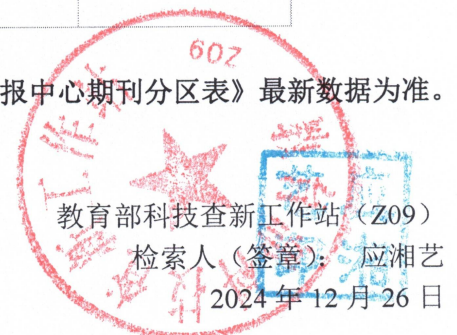
期刊《KNOWLEDGE-BASED SYSTEMS》2023年升级版的中科院期刊分区情况为:

刊名	KNOWLEDGE-BASED SYSTEMS		
年份	2023		
ISSN	0950-7051		
	学科	分区	Top 期刊
大类	计算机科学	1	是
小类	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE 计算机: 人工智能	2	-

注: 1. 期刊影响因子及分区情况最新数据以JCR数据库、《中国科学院文献情报中心期刊分区表》最新数据为准。

2. 以上检索结果来自CALIS查收查引系统。

3. 以上检索结果均得到委托人及被检索作者的确认。



PatchCLIP: Enhancing CLIP with Pair-Images for Zero-Shot Anomaly Detection

Publisher: **IEEE**

[Cite This](#)

 PDF

Yuhu Bai ; Yuhang Dong ; Hongyu Ma ; Binchao Yu ; Zhaofeng Chen ; Longhua Ma **All Authors**

8
Full
Text Views



Abstract

Document Sections

- I. Introduction
- II. Related Work
- III. Method
- IV. Experiment
- V. Conclusion

Authors

Figures

References

Keywords

Metrics

More Like This

Abstract:

The vision-language model CLIP has profoundly transformed the field of zero-shot anomaly detection. Recent studies acquire anomaly maps by aligning images with normal and abnormal prompts. However, CLIP does not effectively interpret text prompts in the context of anomaly detection. As a result, relying solely on text prompts is insufficient, making visual cues particularly crucial. In this paper, we propose PatchClip for zero-shot anomaly detection, integrating both text-image feature matching and image-image matching. We concurrently process two images, treating each as the visual prompt for the other, and combine their respective text-image matching scores. We conduct extensive experiments on anomaly detection benchmarks, demonstrating superior performance compared to previous methods. The code will be open soon.

Published in: [2024 China Automation Congress \(CAC\)](#)

Date of Conference: 01-03 November 2024

DOI: [10.1109/CAC63892.2024.10865319](https://doi.org/10.1109/CAC63892.2024.10865319)

Date Added to IEEE Xplore: 13 February 2025

Publisher: IEEE

► **ISBN Information:**

Conference Location: Qingdao, China

▼ **ISSN Information:**

▼ **Funding Agency:**

SECTION I. Introduction

Industrial image anomaly detection is a critical technology within smart ecosystems, playing a significant role in enhancing system reliability, improving operational efficiency, and supporting data-driven intelligent decisions. By integrating image anomaly detection across various applications, smart ecosystems can more effectively monitor, manage, and optimize their operations. The challenging task involves identifying whether an image is anomalous and segmenting anomalous pixels. The main challenges of anomaly detection arise from the inherent uncertainty and variability of anomaly patterns [1], [2].

Prior anomaly detection methods typically follow an unsupervised paradigm, training solely on a large set of normal images [3]– [9]. However, traditional anomaly detection methods often require extensive data and retraining for each new category or condition, which can be time-consuming and resource-intensive. Therefore, zero-shot anomaly detection (ZSAD) becomes crucial in such scenarios. Zero-shot anomaly detection allows for the rapid deployment of models in dynamic environments where anomalies may evolve or new types may appear. As technology continues to advance, ZSAD methods will likely become increasingly integral to maintaining and optimizing systems across smart ecosystems.

