

样品前处理、重复次数和损耗控制方面存在差异,未来可通过多中心实证研究进一步优化模型参数,提高适用精度。

8 结语

中成药制剂成品抽样量的形成并非依赖经验判断,而应建立在法定检验项目耗样量逐项统计的基础上。按照“项目累加—复核放大—包装换算—损耗修正”的思路,可以构建较为稳定的抽样量计算模型。该模型既能够解释不同剂型、不同品种间抽样量差异,也有助于推动监督抽样和检验实践的标准化、规范化与可复核化。

参考文献

- [1] 国家药典委员会. 中华人民共和国药典: 一部[S]. 北京: 中国医药科技出版社, 2025.
- [2] 国家药典委员会. 中华人民共和国药典: 四部[S]. 北京: 中国医药科技出版社, 2025.
- [3] 王芳, 李强, 赵敏. 药品监督抽样技术规范研究进展[J]. 中国药事, 2022, 36(8): 912-918.
- [4] 张磊, 马会敏. 中成药质量标准与抽样检验关联性分析[J]. 中国中药杂志, 2023, 48(15): 4210-4215.
- [5] 刘洋, 陈晨. 药品检验中样品量确定原则及应用[J]. 药物分析杂志, 2021, 41(10): 1801-1806.
- [6] 孙静, 周楠. 中成药不同剂型抽样检验关键问题探讨[J]. 中国药品标准, 2020, 21(6): 457-462.
- [7] 国家市场监督管理总局. 药品质量抽查检验管理办法[Z]. 北京: 国家市场监督管理总局, 2023.
- [8] World Health Organization. Guidelines for sampling of pharmaceutical products and related materials[R]. Geneva: World Health Organization, 2021.

Research on lightweight detection methods for computer vision in complex industrial scenarios

Tao Song Dingyuan Mo

College of Mathematical and Statistical Sciences Baise University, Baise, Guangxi, 533099, China

Abstract

Complex industrial scenes present challenges such as variable illumination, cluttered backgrounds, and small targets, making it difficult for traditional computer vision detection models to balance detection accuracy and lightweight deployment requirements. This paper reviews four types of lightweight detection methods based on network architecture optimization, model compression, transfer and few-shot learning, and multi-modal fusion. It compares the advantages and disadvantages of various methods, analyzes key issues and solutions in combination with application practices, and verifies the effectiveness of the methods through the construction of an evaluation index system and comparative experiments. The research shows that lightweight methods integrating multiple technologies can better adapt to complex industrial scenes, achieving a balance between accuracy and efficiency, and providing theoretical support for the implementation of intelligent and lightweight industrial detection.

Keywords

Complex industrial scene; Computer vision; Lightweight detection; Multimodal fusion

面向复杂工业场景的计算机视觉轻量化检测方法研究

宋涛 莫定源

百色学院数理科学与统计学院, 中国·广西 百色 533099

摘要

复杂工业场景存在光照多变、背景杂乱、目标微小等问题, 传统计算机视觉检测模型难以兼顾检测精度与轻量化部署需求。本文梳理基于网络架构优化、模型压缩、迁移与小样本学习、多模态融合的四类轻量化检测方法, 对比各类方法优劣, 结合应用实践分析关键问题与解决方案, 通过构建评估指标体系和对比实验验证方法有效性。研究表明, 融合多技术的轻量化方法能更好适配复杂工业场景, 实现精度与效率的平衡, 为工业检测智能化、轻量化落地提供理论支撑。

关键词

复杂工业场景; 计算机视觉; 轻量化检测; 多模态融合

1 引言

在智能制造向纵深发展过程中, 工业检测作为质量控制的核心环节, 对检测实时性、部署便捷性的需求越来越迫切。复杂工业场景中, 由于光照变化、设备干扰、微小故障等因素, 传统检测模型存在精度不足、难以适应新需求的问题, 也有模型烦琐、算力耗费大, 不能部署到边缘设备的痛点。当前, 计算机视觉检测技术虽广泛应用于工业领域, 但轻量化和高精度难以兼得, 仍是影响其落地的关键瓶颈。现有轻量化方法各有局限, 很难充分适配复杂工业场景的多样化需求。本文系统分析各类轻量化检测方法, 结合应用实践和实验验证, 探索适应复杂工业场景的最佳技术路径, 解决精度与轻量化的平衡难题, 助力工业生产向高效、智能、低

成本模式转型。

2 面向复杂工业场景的计算机视觉轻量化检测方法概述

2.1 基于网络架构优化的轻量化检测方法

基于网络架构优化的轻量化检测方法, 核心是重构网络结构、对冗余模块进行简化, 在不会显著造成检测精度下降的状况下, 减少模型参数和计算量, 适配复杂工业场景边缘部署的需求。该方法摒弃传统检测网络的冗余卷积层, 采用深度可分离卷积以及分组卷积等轻量化算子, 结合注意力机制确切聚焦工业目标, 同时优化特征融合结构, 解决复杂场景当中光照变化大、背景杂乱造成的特征提取不准确问题。MobileNet、ShuffleNet 等系列模型作为典型代表, 其通过通道重排、bottleneck 结构设计, 大幅度减少算力耗费。该方法的优点是推理速度快、部署成本低, 可用于工业流水线的实时检测, 然而对架构设计经验的要求比较高, 在微小

【作者简介】宋涛(1971—), 男, 中国河北高阳人, 博士, 高级工程师, 从事数据科学研究。

缺陷检测场景下易造成精度的损失^[1]。

2.2 基于模型压缩的轻量化检测方法

基于模型压缩达成的轻量化检测方法，是在已训练好的高精度检测模型基础上，通过裁剪、量化、蒸馏等手段，清除多余的参数，实现模型“瘦身”，兼顾精度和轻量化。在复杂工业场景下，裁剪策略把焦点放在冗余卷积核与全连接层，留存和工业目标检测有关的核心特征提取组件；把浮点型参数转化为整型的量化过程，降低存储以及计算的成本；知识蒸馏借助教师模型（高精度）来指导学生模型（轻量化）开展训练，传承主要检测能力。此方法不用重新设计网络，适配性强，可快速把成熟的模型应用于工业边缘设备，解决传统模型部署所面临的困难问题。其不足是压缩程度难以平衡，过度压缩会造成复杂场景的检测精度下降，而量化过程有很大可能引入噪声，影响微小缺陷被识别的效果。

2.3 基于迁移学习与小样本学习的轻量化检测方法

基于迁移学习与小样本学习的轻量化检测方法，针对复杂工业场景中缺陷样本不多、标注成本高的痛点，把通用场景训练的模型知识迁移到工业检测任务，结合小样本学习对模型加以优化，实现轻量化和高精度的平衡。这种方法先在大规模通用数据集上来预训练轻量化基础模型，然后运用少量工业缺陷样本对模型参数微调，适合特定的工业场景。以迁移学习降低模型训练的难度，减少参数冗余；借助小样本学习算法，提升模型对稀缺样本的适配能力。其优势是不用大量标注过的样本，训练效率高，适配多种不同的复杂工业场景，但在迁移过程中易有领域适配不足的问题，需针对性优化特征迁移策略。

2.4 基于多模态融合的轻量化检测方法

基于多模态融合的轻量化检测方法，以融合图像、红外、超声等多种工业检测模态特征的方式，弥补单一模态于复杂场景下检测的欠缺，同时采用轻量化的融合组织，避免模型复杂度大幅提升。在复杂工业场景中，单一视觉模态易受光照、遮挡方面的影响，而多模态融合可集合不同模态的优势。以轻量化融合模块为途径，做到多源信息的有效使用。此方法能提高复杂场景下缺陷检测的准确率和鲁棒性，满足工业流水线多方面检测需求。其不足是多模态数据采集成本比较高，融合模块开展轻量化设计难度大，易产生模态冗余以及计算量增加的问题，需优化融合策略，以平衡性能和轻量化的相关需求。

2.5 各类方法的对比分析

四类轻量化检测方法各有优劣，适配不同复杂工业场景的需求。网络架构优化方法推理速度最快、部署成本最低，可用于对实时性要求较高的流水线检测，但架构设计难度大，微小缺陷检测精度一般；模型压缩方法适配性良好，可较快复用成熟的模型，适合已有高精度模型的工业领域，却存在精度与压缩程度难以平衡的问题^[2]。迁移与小样本学习方法无需大量标注样本，适用于缺陷样本不多的场景，但

领域迁移的适配性不足；多模态融合方法有着最强的检测鲁棒性，适宜复杂光照、多维度检测场景，却存在数据采集成本高、融合模块设计复杂的短板。综合来看，单一方法很难完整适配复杂工业场景，实际应用中需结合场景需要，融合多种方法的优势，实现最优检测效果。

3 面向复杂工业场景的轻量化检测方法应用实践

3.1 典型复杂工业场景应用案例

选取机械制造、冶金冶炼、化工管道三个典型复杂工业场景，开展轻量化检测方法并验证效果。机械制造场景中，采用基于网络架构优化的 MobileNet 改进模型，实时对精密零件表面的划痕、变形等缺陷进行检测，部署于生产线边缘设备，检测的速度达到 30 帧 / 秒，缺陷识别准确率 98.2%，较传统模型算力消耗降低 65%。冶金冶炼场景，融合红外与视觉模态的轻量化模型，适配高温环境以及粉尘干扰，做到钢水液位、炉壁破损检测，解决单模态检测精度不精准的问题，响应时间控制在 50ms 内。化工管道场景，采用模型压缩后的 YOLO 轻量化模型，检测管道腐蚀、泄漏缺陷，部署于便携式检测设备，适合野外复杂的环境，检测效率与人工检测相比提升 8 倍，大幅度降低人工检测的风险及成本。

3.2 应用过程中的关键问题与解决方案

应用过程中主要面临三大关键问题，针对性提出解决方案。一是复杂场景（光照、粉尘、振动）的干扰造成检测精度下降，借助自适应光照补偿算法对输入图像加以优化，结合注意力机制强化对目标特征的提取，减少环境干扰影响，提升精度。二是边缘设备算力有限，轻量化模型仍存在推理延迟，通过把模型量化到 INT8 格式，去除冗余特征层，进一步缩小模型体积，保证推理延迟控制在工业所要求的范围内。三是多场景的适配性不佳，同一模型难以适配不同工业场景，设计模块化的模型结构，按照不同场景对特征提取模块和检测头（Head）进行微调，提升模型适应性，降低场景更替造成的部署成本，实现多场景快速推进。

3.3 应用案例总结与启示

综合三个典型场景应用实践总结：轻量化检测方法可有效破解复杂工业场景中“高精度与轻量化难以平衡”“部署困难”等痛点，不同方法适配不同场景所提出的需求，网络架构优化适合实时性要求高的流水线，多模态融合适合干扰强的场景，模型压缩适合快速复用成熟模型。核心启示为，轻量化检测方法落地要与场景特点相结合，杜绝“一刀切”，优先考虑环境干扰、设备算力、样本情况等实际因素^[3]。重视模型的模块化以及可扩展性，方便后续对场景做适配与优化。此外，工业场景落地需有数据积累以及算法迭代，要根据实际检测反馈不断微调模型，才能取得轻量化与检测性能的最优平衡，切实带动工业智能化升级。