

基于深度学习的工业零部件表面缺陷智能检测算法优化

刘兵

浙江煦美智能科技有限公司, 浙江杭州, 310000;

摘要: 在智能制造高速发展的背景下, 工业零部件的表面质量直接关系到产品性能与生产安全。传统的表面缺陷检测方法依赖人工检验或规则化图像处理手段, 存在检测精度低、适应性差与效率瓶颈等问题。随着深度学习技术在计算机视觉领域的广泛应用, 基于深度神经网络的表面缺陷识别逐渐成为智能检测系统的核心。然而, 复杂背景干扰、缺陷特征多样、样本数据不平衡等问题仍制约着算法在实际工业场景中的性能。本文围绕深度学习在工业缺陷检测中的应用展开研究, 从检测模型结构优化、数据增强与小样本学习策略、推理加速与部署适配三个方面系统分析关键技术路径, 并结合典型案例分析其效果表现。研究表明, 通过针对性地优化深度学习模型结构与数据处理流程, 能够显著提升缺陷检测的准确率、实时性与系统鲁棒性, 为高效可靠的工业质检系统提供技术支撑。

关键词: 深度学习; 表面缺陷检测; 工业零部件; 目标识别; 算法优化; 智能质检

DOI: 10. 64216/3080-1508. 25. 02. 017

1 表面缺陷检测中深度学习模型的结构优化

1.1 卷积神经网络在缺陷特征提取中的适应性分析

卷积神经网络 (CNN) 在深度学习领域属于极具代表性的图像处理结构, 它有着局部感知、参数共享以及层级抽象等诸多优点, 所以在表面缺陷识别任务当中得到了极为广泛的应用。工业零部件图像里的缺陷常常呈现出微小且不规则的区域、纹理出现破损或者对比度发生变化等特点, 而 CNN 能够凭借多层卷积这一方式逐步捕捉到从边缘一直到语义级别这样不同尺度的特征。不过, 标准的 CNN 架构要是遇到不同类型的零部件或者复杂材质的情况, 它的特征表达能力是有可能出现退化现象的, 特别是在缺陷区域面积偏小的时候, 就很容易被背景纹理给掩盖住了。为了能够提升对那些微小缺陷的敏感程度, 研究者们一般会在基础的 CNN 结构之上引入诸如空洞卷积、金字塔池化、注意力机制等相关模块, 从而让模型具备更为出色的上下文感知能力以及对局部特征进行强化的能力。与此同时, 为了能够同时兼顾到模型的规模以及训练的效率, 像 MobileNet、EfficientNet 这类轻量化的结构也逐渐被引入到缺陷检测这个领域当中, 它们在减少参数数量的同时还能够保持着比较强的表达能力。

1.2 检测框架结构的选型与改进

当下在目标检测这块, 深度学习框架大致能分成两阶段检测器以及一阶段检测器这两类。像 Faster

R-CNN 就属于前者, 它有着比较高的精度, 不过计算方面的开销也挺大, 所以更适用于那种对精度要求很高的场景。而 YOLO、SSD 这类属于后者, 它们的结构相对来说更为简单明了, 适合在对检测速度有着较高要求的工业场合去应用。在针对工业零部件表面缺陷进行检测的时候, 往往既要保障检测的速度, 又得保证识别的精度, 所以研究者们一直在模型结构的选型上不断去努力平衡、优化。比如说, 可以把 YOLOv5 的主干网络和注意力机制相互结合起来, 引入像 CBAM 或者 SE 这样的模块, 以此来增强模型对于关键区域的响应能力, 这样就能有效地提升它对那些小缺陷的识别能力了。要是碰到那种以精度优先的任务, 那么还能针对 Faster R-CNN 里的 RPN 模块去做一些多尺度锚框的优化处理, 好让它能够适应不同尺寸缺陷进行检测的需求。

除此之外, 还有一部分研究是依据 Transformer 结构来构建表面缺陷检测模型的, 利用它的全球注意力机制来提升模型对图像全局信息的感知能力, 从而突破卷积结构在对远程依赖进行建模时所存在的一些局限。通过对检测框架结构做出有针对性的改造, 并且进行合适的模块组合, 深度学习模型就能够更好地去匹配实际工业环境当中对于检测速度、精度以及模型轻量化等多方面的要求了。

1.3 模型泛化能力与稳定性提升机制

在工业现场实际应用的时候, 模型所具备的泛化能力以及稳定性那可是相当关键的。一方面, 工业现场的零部件种类那叫一个繁杂多样, 而且不同批次之

间的差异也是极为显著的,在这种情况下,检测模型就得有能够去适应那些未曾见过的样本的能力呀。另一方面,环境条件老是在变,像光照情况、反射状况、还有背景干扰这些因素,都会使得模型的输出出现波动。所以,研究者们就在模型设计以及训练这个阶段引入了好多项增强策略,目的就是要提升模型的泛化能力和稳定性。在这些策略当中,运用正则化技术(就好比 Dropout、L2 正则化这类的)来防止模型出现过拟合的情况,这可以说是一种比较基础的做法了。而那些更为先进的方法,就包含了领域自适应训练、多任务联合学习以及不确定性建模等。比如说,通过引入 Domain-Adversarial Training 这种方式,就能让模型学到那种模态不变的特征,这样在不同的数据分布之下,模型也能保持性能的一致性。要是联合分类与定位任务来对模型进行训练的话,那就能够提升整体特征的一致性以及可转移性。还有不确定性建模,就像 Monte Carlo Dropout,它在推理阶段是可以对模型输出的置信度进行评估的,如此一来,也就能够给缺陷判断提供相应的参考依据了。

2 数据驱动优化:增强策略与小样本学习方法

2.1 工业缺陷数据的获取难点与增强需求

和自然图像任务有所不同的是,工业零部件表面缺陷方面的数据存在着采集起来难度颇高、缺陷分布较为稀疏、标签成本也比较高这样的一些特点。大多数的缺陷样本往往是通过人工巡检去做记录或者采用特定的缺陷模拟方式才能够获得,其数量不但有限,而且种类也不均衡,如此一来便使得训练集中正负样本的比例出现了严重失衡的状况,这很容易就致使模型产生偏置的情况。所以呢,数据增强就变成了用来弥补样本不足、扩展数据分布范围的一个重要手段了。常规的一些增强方式,像是旋转、缩放、翻转以及对对比度调整等这些方法,虽说能够增加样本的数量,可却难以模拟出真实缺陷的生成机制。基于此,研究者们便提出了一系列专门针对缺陷图像的增强方法,其中包含了基于图像修补的合成缺陷注入这种方式,还有基于 GAN(生成对抗网络)的缺陷样本生成方法,以及多源数据融合增强策略等。

2.2 小样本学习在稀有缺陷识别中的应用

在实际的工业场景当中,存在着一些稀有且颇为关键的缺陷类型,它们出现的频率是极低的,想要依靠传统的监督学习来完成建模工作,着实存在很大难度。而小样本学习(Few-shot Learning)便由此成

为了能够提升检测系统效能、拓展其检测范围的重要方法。这种方法将元学习的思想引入其中,能让模型在样本数量极少的情形下,较为迅速地去适应新的类别,所以在缺陷样本不够充足的各类任务里面,得到了广泛的应用。其典型的实现形式包含原型网络(Prototypical Network)、匹配网络(Matching Network)以及关系网络(Relation Network)等等。这些模型会去构建相似性度量函数,从而在支持集和查询样本之间构建起类间的关联。此外,把迁移学习的策略与之相结合的做法,同样也被广泛运用起来了。比如说,可以利用已经存在的大规模工业图像预训练模型来实现特征的迁移,并且通过微调的方式,使其能够适配具体的缺陷类别。

2.3 数据不平衡问题的调节与损失函数设计

在实际生产当中,工业缺陷属于异常状态,其样本数量相较于正常样本而言要少得多,如此严重的数据不平衡状况使得在训练过程里,模型会更偏向于去学习正常类别的相关特征,进而降低了对缺陷的敏感性。为了缓解该问题,研究者们提出了诸多调节机制,像样本重加权、过采样、欠采样、在线难例挖掘以及损失函数重构等这些不同的方式。在这之中,Focal Loss 作为处理分类不平衡方面颇具代表性的损失函数,它通过给难分类样本增加一定的权重,切实有效地提升了模型针对少数类样本的识别能力;而 Dice Loss、IoU Loss 等这类损失函数,则是更契合于分割类缺陷检测,能够强化模型对边界特征的呈现与表达。除此之外,那种结合了自适应采样策略的在线学习机制同样展现出了不错的性能,其能够在模型训练的过程之中自动去聚焦那些难样本以及低置信度的区域,以此来提升整体的学习成效。通过对训练数据分布展开较为细致的建模以及对损失函数进行合理的设计安排,检测模型在面对现实当中复杂的数据条件之时,便能够展现出更为强劲的鲁棒性以及更为精准的检测能力。

3 部署优化与系统集成策略

3.1 模型轻量化与推理加速机制

工业零部件表面缺陷检测系统于实际应用之时,往往会被部署在边缘设备里,或者是生产线终端处,又或者是嵌入式系统当中。如此一来,对于模型体积、计算速度以及内存消耗等方面,就有了颇为严格的限制条件。要想确保检测系统具备实时性,同时也拥有可部署性,那么让模型变得轻量化,便成了在部署阶

段极为重要的一项优化任务。通常所采用的策略有模型剪枝、权重量化、蒸馏压缩,还有结构重构等。就拿模型剪枝来说,它是通过移除那些冗余的连接以及通道,在能够保持核心信息表达能力的基础之上,极大地减少模型参数。而量化方法,则是把浮点权重映射成为低比特的表示形式,像是 INT8 或者 INT4 之类的,这不但能够对存储空间进行压缩,而且还可以加快推理的速度。蒸馏方法是借助教师—学生网络机制,把大模型所蕴含的知识迁移到小模型里面去,以此来实现既压缩又兼顾精度的效果。至于结构重构,比如运用 GhostNet、ShuffleNet 等这类轻型结构,是直接从设计层面就降低了网络的复杂度。这些优化技术和深度学习推理引擎,像是 TensorRT、OpenVINO 或者 TFLite 等相结合之后,可以在不同的硬件平台之上达成高效的推理,能够非常显著地提升检测系统在现场部署过程当中运行效率以及响应速度。

3.2 工业检测系统的集成与可视化平台构建

表面缺陷检测算法要实现工程化落地,这可不单取决于模型自身的准确性,还得依靠稳定的系统架构以及良好的人机交互界面才行。在集成这个层面来讲,检测模型得和前端采集设备、后端数据库还有控制系统达成高效的联动,进而形成一套完整的检测闭环流程。常见的架构往往会采用边缘采集、本地识别、结果上传、实时反馈这样的流程设计方式,以此来确保系统能有低延迟的响应,并且实现数据的闭环。而在平台层面呢,去构建可视化监控界面已经变成了提升系统实用性的一个重要办法。借助数据面板、缺陷热图、异常告警、质量趋势分析等这些模块,操作人员便能够实时去监测生产质量的状态,从而及时察觉到潜在的问题。

3.3 系统稳定性验证与部署环境适应性分析

把深度学习模型应用到实际工业环境当中的时候,在部署之前务必要对系统的稳定性还有适应性展开极为严格的测试以及验证工作。首先得针对模型在好多批次、不同批量产品之上的稳定性去做验证,这里面涉及到对准确率、误报率、漏检率等非常关键的指标展开统计分析,目的就是要保证系统能够长时间可靠运行。其次,要开展环境适应性测试,比如在不一样的光照条件下、存在震动干扰的情况下、温湿度出现变化或者工况进行切换的时候,对系统的响应能力加以评估,从而排查出潜在的故障风险。除此之外,在进行系统部署的时候,还得把网络稳定性、电源保

障、硬件兼容性以及维护便捷性等一系列工程方面的因素都考虑进去。在部分高端工业环境里,还得满足行业安全标准才行,就像要通过 EMC 抗干扰、电气隔离、防爆防尘等相关的认证要求。通过去建立一整套标准化的测试与验收流程,便能够最大程度地降低在部署之后系统失效的这种可能性,进而确保检测系统就算是处在复杂多变的工业场景之下,也依然可以保持高精度以及高可靠性地运行。

4 结语

深度学习作为当今人工智能技术的核心力量,正日益成为推动工业智能检测系统高质量发展的关键驱动力。本文聚焦于工业零部件表面缺陷检测任务,系统探讨了深度学习模型的结构优化、数据处理策略与系统部署路径。通过卷积结构增强、检测框架改进与特征泛化机制提升模型表达力,通过数据增强、小样本学习与损失函数调整增强模型对工业复杂样本的适应能力,并通过推理加速、平台集成与部署测试确保系统的实际可用性与工程落地性。研究表明,深度学习与工业场景需求之间的耦合度正持续增强,但仍面临诸如小样本覆盖、实时性优化、系统稳定性保障等关键挑战。

参考文献

- [1] 林灏昶,秦云川,蔡宇辉,等.基于目标检测的图形用户界面控件识别方法[J].南京大学学报(自然科学).2022,58(6).DOI:10.13232/j.cnki.jnju.2022.06.009.
- [2] 董林鹭,向洋,林国军,等.一种基于椒盐噪声密度的滤波算法[J].四川师范大学学报(自然科学版).2021,(2).DOI:10.3969/j.issn.1001-8395.2021.02.020.
- [3] 刘向明,朱启茂,王伟,等.连续铸轧生产线热轧带钢表面缺陷产生原因及改进措施[J].理化检验(物理分册).2021,(7).DOI:10.11973/lh-jy-wl202107018.
- [4] 李少波,杨静,王铮,等.缺陷检测技术的发展与应用研究综述[J].自动化学报.2020,(11).DOI:10.16383/j.aas.c180538.
- [5] 祝安定,施鹏汉,肖宗莹.基于阈值增强的均值滤波[J].电脑知识与技术.2020,(7).
- [6] 黄凤荣,李杨,郭兰申,等.基于 Faster R-CNN 的零件表面缺陷检测算法[J].计算机辅助设计与图形学学报.2020,(6).DOI:10.3724/SP.J.1089.2020.17981.