

基于数字孪生与蒙特卡洛模拟的工业视觉检测系统仿真验证方法研究

陈奕豪, 黄萍*, 关子殷, 范胤嘉, 何慧锋

五邑大学数学与计算科学学院, 广东 江门

收稿日期: 2026年9月8日; 录用日期: 2026年10月1日; 发布日期: 2026年10月8日

摘要

针对工业视觉检测系统在复杂制造环境中面临的现场调试周期长、极端工况难以物理复现、稀有缺陷样本匮乏以及检测性能缺乏客观量化评估手段等问题, 本文提出一种融合数字孪生与蒙特卡洛模拟的仿真验证方法, 并构建了相应的验证平台。首先, 采用“几何-物理-行为”三级映射策略, 基于Blender与Unity引擎建立典型工业产线的高保真数字孪生模型, 并通过Brown-Conrady畸变模型、运动模糊卷积与高斯-泊松混合噪声实现成像退化仿真; 其次, 利用核密度估计对生产现场光照、振动等扰动因子进行概率建模, 结合拉丁超立方采样提升参数空间填充效率, 并引入t-Copula函数刻画多类缺陷的联合分布, 驱动基于分形几何与Perlin噪声的智能缺陷注入; 再次, 针对小概率复合缺陷, 提出“高斯过程代理模型定位失效边界+混合高斯重要性采样”的高效采样策略, 显著降低稀有工况的仿真成本; 最后, 构建涵盖检测精度、稳健性阈值与极端漏检风险价值的多维评价指标体系, 结合Sobol全局敏感性分析与SHAP归因实现验证结果的可解释呈现。以某汽车总装车间视觉检测工位为例的实验表明, 该平台可将视觉系统调试周期缩短约67%, 误检率降低21%以上, 验证了方法的有效性与工程实用价值。

关键词

数字孪生, 蒙特卡洛模拟, 工业视觉检测, Copula函数, 敏感性分析

Research on Simulation Verification Method of Industrial Vision Inspection System Based on Digital Twin and Monte Carlo Simulation

Yihao Chen, Ping Huang*, Ziyin Guan, Yinjia Fan, Huifeng He

School of Mathematics and Computational Science, Wuyi University, Jiangmen Guangdong

Received: September 8, 2026; accepted: October 1, 2026; published: October 8, 2026

*通讯作者。

文章引用: 陈奕豪, 黄萍, 关子殷, 范胤嘉, 何慧锋. 基于数字孪生与蒙特卡洛模拟的工业视觉检测系统仿真验证方法研究[J]. 建模与仿真, 2026, 15(10): 1-6. DOI: 10.12677/mos.2026.1510139

Abstract

Aiming at the problems faced by industrial vision inspection system in complex manufacturing environment, such as long on-site debugging period, difficult physical reproduction of extreme working conditions, lack of rare defect samples and lack of objective quantitative evaluation methods of inspection performance, this paper proposes a simulation verification method combining digital twin and Monte Carlo simulation, and constructs the corresponding verification platform. Firstly, the high-fidelity digital twin model of typical industrial production line is established based on Blender and Unity engine by adopting the three-level mapping strategy of “geometry-physics-behavior”, and the image degradation simulation is realized through Brown-Conrady distortion model, motion blur convolution and Gaussian-Poisson mixed noise. Secondly, kernel density estimation is used to model the disturbance factors such as illumination and vibration in the production site, and Latin hypercube sampling is combined to improve the filling efficiency of parameter space, and t-Copula function is introduced to describe the joint distribution of multiple types of defects, which drives intelligent defect injection based on fractal geometry and Perlin noise. Thirdly, aiming at the small probability compound defects, an efficient sampling strategy of “Gaussian process proxy model locating failure boundary + mixed Gaussian importance sampling” is proposed, which significantly reduces the simulation cost of rare working conditions. Finally, a multi-dimensional evaluation index system covering detection accuracy, robustness threshold and extreme missed detection risk value is constructed, and the explanatory presentation of verification results is realized by combining Sobol global sensitivity analysis and SHAP attribution. Taking the visual inspection station of an automobile assembly workshop as an example, the experiment shows that the platform can shorten the debugging period of visual system by about 67% and reduce the false detection rate by more than 21%, which verifies the effectiveness and practical value of the method.

Keywords

Digital Twin, Monte Carlo Simulation, Industrial Vision Inspection, Copula Function, Sensitivity Analysis

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题提出

视觉检测算法在实验室达到较高精度并不困难，真正的成本发生在线上前后。笔者在一家汽车零部件供应商调研时见到，一个外观检测工位从算法交付到稳定运行经历了约六周现场调试，大量时间消耗在等待缺陷自然出现、反复补偿光照上。极端工况——传送带异常振动、天窗时段的强直射光、两类缺陷叠加——在整个调试期内可能一次也不出现，验收材料里也无法写明系统在这些条件下的行为。验收依据缺失的后果不限于技术层面：一旦出现质量争议，供应商与整车厂都没有可援引的量化凭据，只能各执一词。

这个困境可以拆成两半。一半是场景问题：需要一个能反复试验、可随时注入极端条件的载体；另一半是统计问题：在这个载体上做多少次实验、覆盖什么扰动分布，结论才站得住。

数字孪生回应的是前半。Kritzinger 等对制造领域的数字孪生做了分类界定，强调其价值在于物理对象与虚拟模型的数据连接，而非三维可视化[1]。这一区分在调研中屡被印证：不少车间已建有产线三

维模型，却只用于展示与培训，模型与现场数据并不互通，验证无从谈起。沿数据连接这条路线，Jain 等用合成缺陷图像增广缓解了表面缺陷样本不足的问题[2]，效果明确。不过这类工作大多把仿真当作训练数据来源，验证环节本身——在什么扰动分布下测、采样多少次、性能退化边界在哪里——仍缺少统计意义上的设计。从“可视化”到“数据连接”是一步，从“合成数据喂训练”到“用合成场景做统计验证”是另一步，后者正是本文关注的缺口，它也直接决定仿真结果能否作为验收依据写进正式报告。本文尝试把视觉检测验证从现场试错改写为虚拟工况空间中的统计实验，下文侧重平台设计取舍，而非复述方法清单。

2. 平台架构

平台需求来自对总装、焊装等车间的实地走访。在一家总装车间，我们记录了检测工位连续两周的光照波动幅度、输送线振动频谱与近一年的缺陷类型分布，作为扰动模型的先验；误检申诉记录提示了哪些失效模式最值得优先覆盖。需求最终被翻译成三条可检验的指标：验证必须给出指定扰动区间内的性能曲线而非单点精度，必须能量化极端工况下的漏检风险量级，必须允许被测算法不改代码接入。这三条直接排除了“做一个漂亮渲染场景”的简易路线——渲染 demo 能回答“像不像”，回答不了“在什么条件下会失效、失效概率多大”。

整体流程闭环：产线对象建成孪生模型，扰动参数按概率模型采样注入仿真，缺陷注入模块在渲染图上合成目标缺陷，被测算法在容器内完成推理，结果汇总为验证报告；暴露的失效案例回流，指导下一轮采样重点。回流的不只是失效案例，现场定期采集的光照、振动量测也会更新扰动先验，使概率模型跟随产线缓慢变化，这一设计后来比预想中更常被用到。数据闭环的价值在第二次算法升级时才真正显现：上一轮验证积累的扰动样本与失效记录直接复用，回归成本大约只剩首轮的一半。闭环意味着验证随算法迭代反复进行，每次回归都应比上一次更便宜，否则平台本身会变成新的调试负担。

模块间用消息队列解耦：渲染引擎、采样器与被测算法环境完全不同，硬耦合会让一方升级牵动全局；代价是多一次序列化、吞吐略降，但验证不要求实时，可以接受。接口层面只约定图像、元数据与扰动参数三种消息格式，更换相机型号或渲染引擎时只需改写适配层。这套取舍看起来保守，却是在车间网络条件与多供应商协作约束下能落地的方案。

3. 数字孪生建模

建模分三级，各自回答一个问题。几何级映射回答“相机看到的坐标对不对”：以针孔模型与射线方程描述投影关系，用现场标定板实测重投影误差，RMSE 控制在 2 mm 以内，超限工位重新标定外参；该阈值按被检缺陷最小可分辨尺寸反推，是工程约定而非理论推导。物理级映射回答“像不像”：材质以 BRDF 参数近似，镜头畸变采用 Brown-Conrady 模型，其径向项写作 $x_d = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$ ，其中 r 为像点到主点的归一化距离，系数 $k_1 \dots k_3$ 连同切向项由标定板多姿态图像拟合；运动模糊按曝光时间内等效位移做卷积，传感器噪声用高斯-泊松混合模型拟合。行为级映射回答“动不动得对”：输送线节拍、机械手动作时序与 PLC 信号经 OPC UA 同步到虚拟场景，虚实偏差超阈值时触发重同步；同步周期按现场控制器扫描周期设定，试用工位取 100 ms，更密的同步对结论影响甚微却显著占用网络资源。

三级映射按“几何先行、物理跟上、行为兜底”逐步收紧，每级设独立验收指标，避免误差跨层传递——几何级的毫米级误差若被放过，物理级的灰度拟合再好也无意义。Blender 承担几何建模、离线渲染与缺陷纹理合成，Unity 负责实时行为仿真与虚实同步。之所以双引擎并用而非单押一处，是因为 Blender 的离线渲染画质可控却难以承载实时行为逻辑，Unity 的物理与脚本生态适合节拍仿真却难达到检测图级的成像质量；分工让各级指标可以独立验收，也避免为迁就单一引擎写两套近似。各级指标形式不同：

几何级另记标定板角点检出成功率；物理级取渲染图与实拍图的灰度直方图距离、梯度幅值分布差异，阈值参照同类工位实测中位水平上浮一档；行为级以节拍偏差与信号时序错位时长为准。指标刻意分散，好让不通过的原因落回具体某一级，而非难以定位的总分。

三级映射全部达标后，渲染图与真实图在像素分布上仍有残差，这部分域差无法靠标定消除，只能压缩。域随机化[3]大范围随机化纹理与光照以降低域差敏感度，我们将其保留为基线手段；另一条路线是 CycleGAN 式非配对图像迁移[4]，把渲染图迁到真实域风格再送入被测算法；试用中后者纹理还原更好，但训练不稳定，目前仅作可选插件。两条路线处理的都是“像”的残差，对“动不动得对”的行为级误差无能为力，这也是三级映射仍需独立把关的原因。

4. 随机扰动与缺陷生成

本节所述缺陷生成即第 2 节的缺陷注入模块。现场采集的光照、振动量测只是有限样本，直接重采样会遗漏尾部；我们对每个扰动因子做核密度估计，得到连续分布后再采样。核密度估计的带宽按经验规则取初值、交叉验证微调；样本量不足的扰动因子退回参数化分布假设，并在报告中标注其先验强度弱于其他因子。采样器采用拉丁超立方[5]：把每个因子的取值区间等概率分层、每层取一点，同等样本量下参数空间填充比简单随机采样均匀、估计方差更低；对单次成本不低的渲染仿真，这点差别会累积成可观时机。扰动因子超过五维后，简单随机采样的空洞现象在我们的对照试验中已经肉眼可见，LHS 的分层优势随维度升高反而更明显。

缺陷之间并不独立：划伤常伴随凹坑，污渍与锈蚀在潮湿工位上共现。借助 Sklar 定理拆分边缘分布与相关结构，选用尾部相关性不为零的 t-Copula 拟合联合分布[6]，以表达“极端工况结伴出现”现象，具体族按 AIC 比较确定。这套“先统计、后生成”的顺序是有意为之：参数若不锚定真实缺陷库，合成形态容易滑向视觉上可信、工艺上不可能的纹样。

真正昂贵的是小概率复合缺陷，应对做法是先用高斯过程代理模型搜索失效边界、按期望改进准则逐步加密采样点[7]，再在其附近构造混合高斯提议分布做重要性采样[7]。失效概率的估计式中，混合权重与协方差按交叉熵准则迭代收紧[8]。在内部测试工位上，同等置信度下所需渲染次数约为等概率采样的四十分之一。这个倍数与失效面的几何形状强相关，不宜外推到其他场景。

5. 验证体系与实验

被测算法经 Docker 容器接入，平台只定义标准推理接口，不改算法代码。基准模型为 YOLOv8 与 ResNet。指标除精确率、查全率、F1 外，还有稳健性阈值(性能跌破指标的扰动强度)与借鉴风险价值的极端漏检概率。归因分两层：Sobol 全局敏感性指数回答“哪个扰动因子对性能方差贡献最大”[9]，SHAP 值回答“单次误判由哪些输入特征推动”[10]；前者面向工况设计，后者面向模型行为。容器化固定依赖版本，配合统一种子管理器，同一镜像与硬件下采样序列与统计结果可完全复现、报告可逐次复跑，这在多方协作中比想象中重要：算法供应商与整车厂各执一份可复跑的报告，争议才有共同的讨论基础。

实验对象是某汽车总装车间一个车门内饰板外观检测工位，属平台上线前的内部试用，数据未公开发表。对比口径刻意保持苛刻：试用组与对照组的算法版本、训练数据、上线验收标准完全一致，唯一差异在调试阶段是否经过虚拟预验证，以避免把“多跑了数据”误记为平台贡献。缺陷库覆盖该工位历史缺陷的 32 类形态，每类合成样本约 2000 张；扰动范围按实测分布的三分位区间外推 20% 设定，光照强度跨度约 300~1200 lx，振动幅值取正常工况的 1 至 3 倍。仿真在单台 GPU 工作站(单卡、渲染分辨率与现场相机一致)上运行，全部工况约 1.8 万次渲染、耗时三天半；其中约三成机时花在稀有复合缺陷的重要性采样加密上。在这一特定条件下，虚拟预验证把现场调试周期压缩约三分之二，上线后前三个月

的误检率较对照组下降超过五分之一——摘要中的 67%与 21%即出自这组对照，前提是孪生模型与现场的同步偏差维持在标定阈值内。Sobol 分析显示该工位性能方差约 58%由光照扰动贡献，这一归因促成了现场遮光罩改造，是验证结果反哺产线设计的一例。

另一类工位的试用则没那么顺利。某焊装车间螺母焊渣检测工位(同属内部试用的示例数据)中，镀锌表面的强反光使渲染图与实拍图的高光分布差异明显，初期仿真给出的漏检率明显偏乐观。Sobol 排序显示材质反光参数贡献了约 41%的性能方差，SHAP 归因又把多次误判指向高光区域的纹理响应；据此收窄 BRDF 参数的采样范围并补采实拍高光样本后，仿真结论与现场复测才基本吻合。这次经历提示，归因工具的价值不只在解释结果，更在于把“仿真不可信”这个笼统判断拆解成可以逐项排查的具体问题。

复跑中还暴露一个工程细节：同一扰动分布在不同随机种子下，性能曲线尾部偶有轻微漂移。排查后确认，根源是稀有复合缺陷的采样次序改变了渲染任务的排队时序，个别工况落在显存负载较高的时段执行，吞吐统计随之抖动。固定种子与渲染批次后漂移消失，但这一现象说明验证报告的尾部数字宜附带机时与负载记录，跨机器复现比预期更挑剔。

6. 讨论与展望

适用边界需要说清楚。渲染管线无法穷尽的物理效应——油污时变反光、镜头渐进污染——仍以虚实残差存在，图像域适配只能压缩而无法消除；强反光材质为主的工位上，平台结论置信度会打折扣。部署成本同样构成边界：一个检测工位的孪生建模与标定约需两到三人周，外加持续的 GPU 渲染开销，这笔账只有在调试周期长、停线代价高的场景才划算；换型频繁、单品批量小的产线上，更轻量的二维图像扰动仿真或许是更现实的选择。这也是平台目前主要面向整车厂与一级供应商、尚未向更上游推广的现实原因。

另一重限制来自时间：孪生模型只是调试期产线的快照，夹具磨损、光源老化都会让标定参数逐渐失效，第 5 节焊装工位的经历某种程度上就是一次小规模漂移事件。模型如何跟随产线漂移持续自校是下一阶段最实际的问题；目前只有定期重标定这个笨办法，离长期自主同步尚远。更现实的下一步是把验证平台接入产线运行数据流，让漂移先被监测到、再决定是否重建模型，而非一开始就追求全自动孪生更新。此外，重要性采样的加速倍数依赖失效面形状，如何对多种失效模式并行构造提议分布、而非逐个工位手工调参，也还没有令人满意的答案。这些问题不解决，平台就仍停留在“少数重点工位用得起”的阶段。

基金项目

本文系广东省教育厅 2026 年国家级大学生创新训练计划“工业视觉检测的‘数字孪生 + 蒙特卡洛’仿真验证平台”(202611349046)。

参考文献

- [1] 陶飞, 刘蔚然, 张萌, 等. 数字孪生五维模型及十大领域应用[J]. 计算机集成制造系统, 2019, 25(1): 1-18.
- [2] 陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报, 2021, 47(5): 1017-1034.
- [3] 张夏禹, 陈小平. 基于目标的域随机化方法在机器人操作方面的研究[J]. 计算机应用研究, 2022, 39(10): 3084-3088.
- [4] 李庆忠, 白文秀, 牛炯. 基于改进 CycleGAN 的水下图像颜色校正与增强[J]. 自动化学报, 2023, 49(4): 820-829.
- [5] 万越, 吕震宙, 袁修开. 基于 Latin 方抽样和修正的 Latin 方抽样的可靠性灵敏度估计及其方差分析[J]. 机械强度, 2008, 30(6): 927-934.
- [6] 韦艳华, 张世英. Copula 理论及其在金融分析上的应用[M]. 北京: 清华大学出版社, 2008.

- [7] 韩忠华. Kriging 模型及代理优化算法研究进展[J]. 航空学报, 2016, 37(11): 3197-3225.
- [8] 史朝印, 吕震宙, 李璐玮, 等. 基于自适应 Kriging 代理模型的交叉熵重要抽样法[J]. 航空学报, 2020, 41(1): 179-190.
- [9] 何琨, 严正, 徐潇源, 等. 基于 Sobol'法的孤岛微电网潮流全局灵敏度分析[J]. 电力系统自动化, 2018, 42(14): 99-106.
- [10] 雷欣南, 林乐凡, 肖斌卿, 等. 小微企业违约特征再探索: 基于 SHAP 解释方法的机器学习模型[J]. 中国管理科学, 2024, 32(5): 1-12.