

团 体 标 准

T/QME 0303-2022

基于深度学习的装配过程监测 技术规范

Assembly process monitoring based on deep learning

-technical specifications

(发布稿)

2022-12-21 发布

2022-12-21 实施

青岛市机械电子工程学会 发 布

目 次

前 言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 系统技术框架 2

5 图像样本集制备要求 3

6 深度学习模型训练和评价 4

7 系统功能要求 4

8 系统性能要求 5

前 言

本文件按照 GB/T 1.1-2020 《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定编写。

请注意本文件的某些内容可能涉及专利，本文件的发布机构不承担识别这些专利的责任。

本文件由青岛理工大学提出。

本文件由青岛市机械电子工程学会归口。

本文件起草单位：青岛理工大学、西安交通大学、上海交通大学、北京邮电大学、青岛鹏海软件有限公司、哈尔滨工业大学、青岛市产品质量检验技术研究所、青岛檬豆网络科技有限公司。

本文件主要起草人：陈成军、李东年、董海韬、代成刚、潘勇、郑帅、武殿梁、辛若杰、方维、翟伟伟、张盛平、赵正旭、洪军、王晓宾、张强、王金磊、郭阳、孙显彬、霍胜军、井陆阳、战卫侠、官源林、张庆海、王爱玲、文妍、赵景波、郑鑫。

本文件为首次制定。

基于深度学习的装配过程监测 技术规范

1 范围

本文件规定了基于深度学习的装配过程监测的术语和定义、技术框架、功能要求、性能要求等。
本文件适用于各种基于深度学习的装配过程监测任务。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 5271.13-2008 信息技术 词汇 第 13 部分：计算机图形

GB/T 5271.24-2000 信息技术 词汇 第 24 部分：计算机集成制造

GB/T 38247-2019 信息技术 增强现实 术语

GB/T 41864-2022 信息技术 计算机视觉 术语

GB/T 41867-2022 信息技术 人工智能 术语

GB/T 40659-2021 智能制造 机器视觉在线检测系统 通用要求

T/CARA 0003-2021 AR 内容格式技术规范

3 术语和定义

GB/T 5271.13-2008、GB/T 5271.24-2000、GB/T 38247-2019、GB/T 41864-2022、GB/T 41867-2022 界定的以及以下术语和定义适用于本文件。

3.1

装配体虚拟模型 Virtual model of assembly

通过 CAD 建模软件和虚拟现实可视化建模软件为待识别的装配体建立的可导入图形渲染引擎进行渲染的三维模型。

3.2

合成图像训练集 Synthetic image training set

利用装配体虚拟模型由图形渲染引擎以离屏渲染方式合成的图像训练集，用于训练深度学习模型。

3.3

真实图像训练集 Real image training set

利用真实相机采集的装配体 RGB 图像和深度图像制备的图像训练集，采用人工标记，用于训练深度学习模型。

3.4

T/QME 0303-2022

合成图像测试集 Synthetic image test set

用于测试深度学习模型的合成图像样本集。

3.5

真实图像测试集 Real image test set

用于测试深度学习模型的真实图像样本集。

3.6

语义分割 Semantic Segmentation

语义分割任务需要对图像中所有像素点进行分类，将相同类别的像素归为相同的标签（常常采用相同的像素点表示）。

3.7

实例分割 Instance Segmentation

实例分割不仅需要将图像中所有像素进行分类，还需要区分相同类别中不同个体。

3.8

目标检测 Object Detection

目标检测需要精确地定位出图像中某一物体类别信息和所在位置。

3.9

变化检测 Change Detection

变化检测是从不同时期采集的图像数据中定量分析和确定场景变化的特征与过程。

3.10

准确率 Accuracy

所有类别正确识别的样本数量占所有样本数量的比值。

3.11

精确率 Precision

正确识别为某类样本的样本数量占所有被识别为该类的样本数量的比值。

3.12

召回率 Recall

正确识别为某类样本的样本数量与该类样本实际数量的比值。

3.13

交并比 Intersection over Union(IoU)

检测结果区域与样本标注区域的交集与并集的比值。

4 系统技术框架

基于深度学习的装配过程监测宜采用语义分割、实例分割、目标检测、变化检测等深度学习模型，

实现装配体图像的检测处理，如图 1 所示，主要包括图像预处理、模型训练和检测判别三个主要步骤。

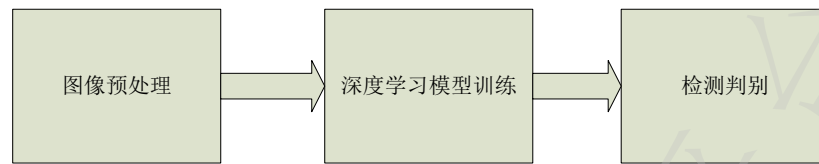


图 1 系统框架

- a) 图像预处理：对被检测装配体的图像样本进行几何变换、归一化、平滑、修复和增强等预处理，优化和改善对装配过程监测有影响的图像质量指标。
- b) 模型训练：将完成图像预处理的装配体样本图像和对应判别结果标签组成训练样本集，输入到语义分割、实例分割、目标检测、变化检测等深度学习模型中，进行参数训练，生成装配过程监测深度学习模型。
- c) 检测判别：将完成图像预处理的检测对象图像输入到训练好的装配过程监测深度学习模型中进行判别，输出判别结果。

5 图像样本集制备要求

5.1 合成图像训练集制备

5.1.1 装配体虚拟模型

- a) 应先采用 CAD 建模软件建立待识别的装配体三维模型；
- b) CAD 建模软件中建立的装配体三维模型应导出为 obj 格式文件，进而导入到虚拟现实可视化建模软件中进行纹理、材质、颜色属性的添加；
- c) 模型名称、网格名称、材质名称、纹理贴图名称应使用英文字母和数字，不应使用中文、空格、标点符号或特殊字符等，不可重名；
- d) 模型制作单位统一为：毫米（mm）；
- e) 坐标系原点应位于装配体的中心或者底部；
- f) 纹理贴图文件格式应符合 T/CARA 0003—2021 6.1.4.1 的规定，文件尺寸不应超过 1024×1024 像素且分辨率应为 2 的幂，不应出现长宽比例不一致贴图，推荐使用 256×256、512×512 像素；
- g) 不应在装配体场景中设置相机节点和灯光节点；
- h) 用于渲染语义分割标签图像的装配体虚拟模型的网格应采用颜色标记，并采用 FLAT 着色模式。

5.1.2 合成图像渲染

- a) 合成训练图像应利用装配体虚拟模型，由图形渲染引擎以离屏渲染方式合成；
- b) 渲染不同视角下的装配体图像时，应在限定的 6 自由度位姿空间中对位姿进行均匀采样；
- c) 合成深度图像应由 GPU 深度缓存 Z-buffer 中储存的数据生成；
- d) 计算机中深度缓存采用伪距离的方式存储深度信息，合成深度图像的渲染应从伪距离换算出距离；
- e) 合成彩色图像应由 GPU 颜色缓存中储存的数据生成；
- f) 虚拟相机的远近裁剪平面、水平垂直视场角等参数设置应与采集真实图像的真实相机的参数相同。

5.2 真实图像训练集制备

- a) 用于采集真实装配体深度图像的深度相机应满足装配过程检测任务的距离测量精度要求和分辨率要求；
- b) 用于采集真实装配体 RGB 图像的彩色相机应满足装配过程检测任务的分辨率要求；
- c) 真实图像训练集的语义分割颜色标签应利用图像处理软件的颜色标记功能进行人工标记；
- d) 真实图像采集过程中，应设置合理的光源照明，以利于真实装配体前景区域的分割提取。

5.3 合成图像测试集制备

合成图像测试集的制备，应满足 5.1.1 的要求，选取不同于合成训练集的图像样本。

5.4 真实图像测试集制备

真实图像测试集的制备，应满足 5.1.2 的要求，选取不同于真实训练集的图像样本。

6 深度学习模型训练和评价

- a) 系统应能够将图像预处理后的带标签的装配体图像样本训练集，输入到语义分割、实例分割、目标检测、变化检测等深度学习模型中进行参数训练，生成装配过程监测数学模型；
- b) 深度学习模型的模型深度与参数规模应与装配体图像训练集的数据量和信息量相符，以避免模型的欠拟合和过拟合；
- c) 训练好的深度学习语义分割、实例分割、目标检测、变化检测等模型宜采用准确率、精确率、召回率、F1、交并比（IoU）、平均交并比（MIOU）、对输入图像的响应速度等指标进行综合评价；
- d) 通过训练和评价的反复迭代，确定最终的深度学习模型。

准确率、精确率、召回率、F1、交并比（IoU）、平均交并比（MIOU）的计算过程如下：

$$\text{准确率: Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$\text{精确率: Precision} = \frac{TP}{TP+FP}$$

$$\text{召回率: Recall} = \frac{TP}{TP+FN}$$

$$\text{F1: F1} = \frac{2TP}{2TP+FP+FN}$$

$$\text{交并比 (IoU): IoU} = \frac{TP}{FP+FN+TP}$$

$$\text{平均交并比 (MIOU): MIOU} = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FP+FN+TP}$$

其中， TP 为正类判定为正类的样本数量， FP 为负类判定为正类的样本数量， TN 为负类判定为正类的样本数量， FN 为正类判定为负类的样本数量， $k+1$ 为类别数量。

对输入图像的响应速度采用每秒处理帧数（FPS）进行衡量。

7 系统功能要求

- a) 系统应能够对装配过程各阶段的装配体零部件进行语义分割和实例分割；
- b) 系统应能够对装配过程中装配体新增零部件进行检测和分割；

- c) 装配过程中各零部件装配结束后系统应能够对零部件的错装和漏装情况进行检测;
- d) 装配过程在线监测的输入和输出模块应满足 GB/T 40659-2021 5.2 和 5.4 的规定。

8 系统性能要求

- a) 装配过程监测系统模型判别结果的准确率、精确率和召回率应满足用户要求;
 - b) 装配过程监测系统对输入图像的响应速度应满足用户要求, 并与资源消耗和检测效果相适应;
 - c) 装配过程监测处理过程应满足硬件负载, 并与图像采集、输入、输出的资源占用相适应;
 - d) 装配过程监测系统应能够适应外部需求的变化;
 - e) 多次重复执行相同的装配监测任务, 判别结果应一致;
 - f) 系统可在装配过程监测各阶段和各种工作环境中对装配体进行稳定的图像采集和分析判别;
 - g) 当装配环境发生异常变化时, 系统可自行学习并检测出异常情况。
-