

AI+CAD 评测系列技术白皮书

通用大模型的 CAD 能力评测报告 ——以 GPT-6 Astra 为例

邹强 沈佳军 杨智杰 王禹皓 冷石天一 尹东硕

浙江大学 DMA 实验室

2026 年 10 月

编写说明

本报告由浙江大学 DMA（Design and Manufacturing Automation Lab）实验室编写。

团队简介

DMA 实验室隶属于浙江大学 CAD&CG 全国重点实验室，主要从事几何计算、智能 CAD 研究及 CAD/CAM 软件研发。

主页（中文）：<https://person.zju.edu.cn/qiangzou>

主页（英文）：<https://qiang-zou.github.io/>

作者贡献

邹强：报告大纲与测试大纲设计、测试结果综合分析、报告撰写及最终审定。

沈佳军：实验实施、报告撰写及统稿。

杨智杰：报告大纲设计、实验实施与报告撰写。

王禹皓、冷石天一、尹东硕：实验实施与报告撰写。

前言

人工智能技术的快速发展正在重塑多个领域的技术范式。在软件开发、科学研究等领域，大模型开始承担原本依赖专门方法与流程的工作，改变既有的人机分工。在 CAD 领域，三维理解与建模能力的不足，曾被视为通用大模型难以深入专业任务的重要障碍。然而，随着 GPT-6 Astra 等新一代大模型与智能体技术的发展，一些此前被认为超出通用大模型能力范围的 CAD 任务，已经出现成功实例。这一变化将如何影响 CAD 软件的技术体系与使用方式，需要系统考察。

对这一变化的认识，需要建立在 CAD 自身的专业要求之上。CAD 模型承载设计意图，支撑后续编辑、分析与制造，其质量直接关系到设计方案能否被准确表达和持续使用。通用模型在具体任务中的表现展示了新的可能，但这些能力能够覆盖多大范围、在何种条件下成立、能否可靠满足工程要求，仍需通过实际任务加以辨明。

为此，本报告以 GPT-6 Astra 为评测对象，围绕 CAD 模型从理解、构建到后续修改的主要环节，组织六类典型任务，包括零件分类、特征识别、工程图到 CAD 模型生成、文本到 CAD 模型生成、图像到 CAD 模型生成，以及 B-rep 模型参数化重建。测试在统一的基本环境下开展，并设计了不同复杂程度的单零件与装配体测试案例。能力判断以实际输出的核验为依据：对识别结果检查其正确性与几何精度，对生成结果检查几何与尺寸、实体有效性和结构关系，对参数化重建结果进一步开展编辑验证。报告结合结果验证与执行过程分析，逐项考察各类任务的完成情况，再综合不同任务中的表现，分析 GPT-6 Astra 可能的推理过程，并总结已经表现出的能力、仍然存在的困难及其可能原因。

本报告面向 CAD 研究者、软件开发者和工程技术人员，旨在为认识通用大模型的 CAD 能力及其局限、审视 CAD 技术体系的变化提供实证依据并结合能力进展与失败案例，重新审视 CAD 研究问题价值。报告反映特定系统、样本与测试条件下的阶段性认识，相关结论仍需在更广泛的任务与实际应用中检验。

摘要

本报告以 GPT-6 Astra（下称 GPT-6）为评测对象，围绕零件分类、特征识别、工程图到 CAD 模型生成、文本到 CAD 模型生成、图像到 CAD 模型生成及 B-rep 模型参数化重建六类任务，考察其在不同复杂度和输入条件下的表现。能力判断以实际结果的专业核验为基础，并结合具体案例与可观察的执行过程展开。

一、核心结论

测评表明，通用大模型的 CAD 能力已开始显现工程应用价值，处于“奇点”阶段。与现有研究中“玩具”数据集和模型不同，GPT-6 在一系列真实场景中取得成功：能够为概念设计提供初步方案，为详细设计提供初始模型。然而，GPT-6 在实体有效性、功能完整性和参数可编辑性仍需专业核验，必要时修正或重建，同时较长的任务执行时间也制约着实际可用性。

（1）通用大模型已经能够承担多类 CAD 任务

GPT-6 能够识别已有 B-rep 模型中的零件类别与几何/加工特征，根据工程图、图像或文字描述生成三维模型，也能在部分任务中依据功能需求选择设计方案、确定尺寸并开展分析校核。本次测试已出现较复杂零件特征识别、中等复杂模型生成与重建，以及较复杂装配体构建的成功案例。

（2）生成模型的实体有效性和功能可行性仍不稳定

生成的 CAD 模型大多在外形上接近目标要求，但内部结构、关键尺寸和特征布局仍可能存在错误。这些问题既出现在单个零件中，也出现在装配体中：孔道可能未按要求连通或隔离，零件之间也可能缺少必要的固定和传动连接。部分错误对外形影响很小，却直接影响关键功能的实现。这类错误不能仅归因于形状复杂或输入信息不足：简单模型同样发现错误，即使提供完整的目标几何信息，也不能保证重建结果正确。

（3）建立三维模型后，仍难保证后续参数化修改正确

对于形状简单、特征数量较少、相互依赖关系简单的模型，GPT-6 已经能够建立可编辑的参数化模型，较好地支持尺寸修改及相关特征的联动。对于中等和复杂模型，所建立的特征依赖关系往往不完整，修改尺寸后可能出现孔口封闭、局部特征消失或零件错位，目前尚不能稳定生成具有良好可编辑性的参数化模型。

（4）GPT-6 已能自主检查和修改模型，但纠错能力仍不稳定

GPT-6 能够在建模过程中检查结果、发现部分问题并反复修改，这正是其完成一些复杂任务的核心原因。然而，测试结果发现，迭代过程中误差曲线经常出现波动，部分修改

反而使误差增大。这表明，GPT-6 对误差来源、修改对象及相应操作的判断仍不够准确。同时，几何特征与参数相互依赖，一处修改可能影响多个结构和设计要求。GPT-6 尚不能充分识别这些关联关系，也难以预判和控制修改的连带影响，因而可能在修正原有问题时引入新的偏差。

（5）建模耗时过长，制约实际可用性

GPT-6 在建模过程中往往需要反复生成、检查和修正模型，带来较大的时间开销。测试中，多类建模与重建任务的平均用时达到数十分钟，部分复杂案例超过 1 小时，最长达 3 小时 41 分钟。较长的执行时间使其难以满足设计过程中快速反馈和频繁修改的需求，尽管建模效果已有进展，当前实际可用性仍然有限。

表 1 GPT-6 在 CAD 任务中的当前能力与主要不足

测评内容	当前能力	主要不足
零件分类	能够结合几何信息与工程知识判断零件类别	细分类型判断仍不够可靠，可能将局部结构的功能误认为零件的整体用途
特征识别	能够识别孔、槽、型腔等加工特征，划分特征实例	复杂区域仍存在实例划分和类型判断错误，特别是对于嵌套、相交等复杂特征
工程图到 CAD 模型重建/生成	能够综合多视图、剖视与尺寸标注，恢复主要外形和内部结构，并完成部分多零件模型的装配	复杂结构层次与关键尺寸仍有误判，典型的有：遗漏局部特征或造成装配干涉
文本到 CAD 模型生成	能够按几何描述构造模型，依据功能需求选择设计方案，并完成建模与分析计算	复杂几何特征、孔道连通、局部过渡、零件固定及传动连接仍经常出错，功能要求与实体有效性尚不能稳定保证
图像到 CAD 模型重建/生成	能够根据多视图图像恢复主体轮廓、主要特征及零件布局	复杂曲面与连接过渡常被简化，实体连接及装配接口仍有偏差，无尺寸参照时难以确定实际尺度
B-rep 模型参数化重建	能够恢复简单和中等复杂模型的主体几何，建立参数化模型并支持部分尺寸修改	复杂曲面与内部孔道等几何的重建存在严重偏差，特征依赖关系经常不完整，修改后相关特征无法正确联动
Agentic 迭代优化	能够利用反馈补齐特征、调整尺寸与位置，降低几何误差，提高要求满足程度	仍难准确确定修改对象与范围，多轮迭代后仍可能遗留无效实体、功能缺失和参数修改失败等问题；迭代耗时较长，限制实际可用性

二、分项测评结论

(1) 零件分类

零件分类任务考察 GPT-6 能否依据三维 CAD 模型的几何与拓扑信息，正确判断零件类别。测试选取 12 个模型，覆盖 10 类零件，简单、中等和复杂样本各 4 个，代表样本见图 1。

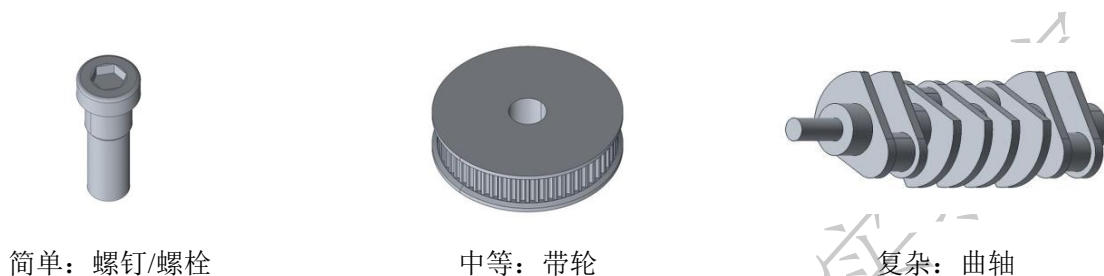
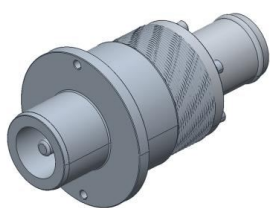


图 1 零件分类任务的不同复杂度样本

测试结果显示，在给定 10 个候选零件类别名称且默认开启网络搜索功能的条件下，GPT-6 对 12 个样本全部分类正确，准确率为 100%。简单、中等和复杂样本的平均用时分别为 2 分 30 秒、2 分钟和 2 分 30 秒，单个样本最长用时为 4 分钟。在不提供零件类别名称或关闭网络搜索功能的条件下，各自选取了 2 个代表性复杂样本进行测试，分类结果仍全部正确。

测试中出现的典型错误是：在不提供零件类别名称且关闭网络搜索的条件下，GPT-6 将管路快速接头误判为止回阀（单向阀），见图 2。GPT-6 虽识别出了内部的弹簧、阀杆和阀瓣，但将阀结构的功能等同于整个接头的用途，导致整体类别误判。



实际零件：管路快速接头
GPT-6 判断：止回阀（单向阀）

图 2 管路快速接头的类别误判

总体而言，GPT-6 目前的能力是：

- 对于典型零件，GPT-6 能够正确识别零件类别，还能指出具体细分类型和功能用途；
- 对于外形不典型或与其他类别相近的零件，GPT-6 能正确识别零件类别，但不一定能判断零件的具体细分类型。
- 对于装配体，GPT-6 可能将局部结构功能误认为整体用途，导致零件类别误判。

（2）特征识别

特征识别任务考察 GPT-6 能否从 B-rep 模型中自主划分加工特征实例、判断类别与子类型，并识别特征之间的相互关系（测试时不预设特征类型的候选列表）。测试选取简单、中等和复杂模型各 2 个。代表样本见图 3。

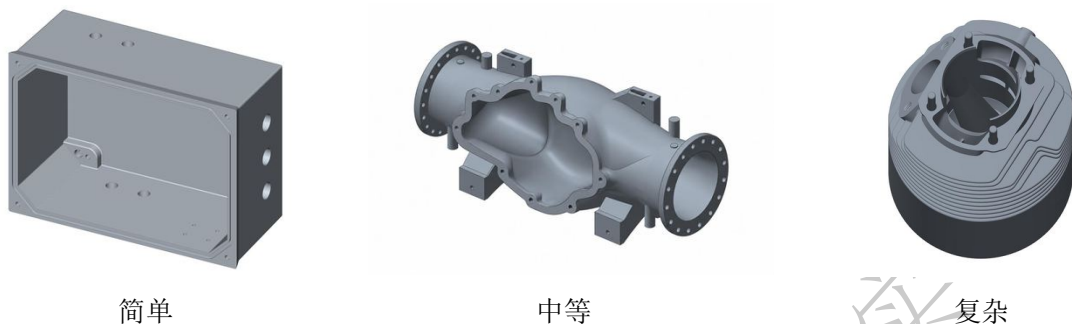
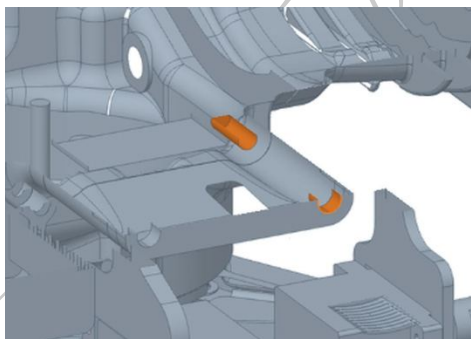


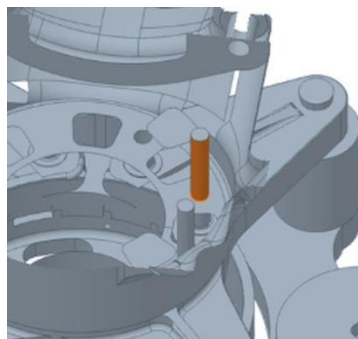
图 3 特征识别任务的不同复杂度样本

测试结果显示，在 6 个测试样例中，3 个样例的特征识别结果正确；2 个样例能正确识别出特征，但还会额外将一些非特征面组也标记为特征；剩下 1 个样例大部分特征识别正确，但存在少量错误。简单、中等和复杂样本的平均用时分别为 30 分 30 秒、40 分 30 秒和 1 小时 25 分钟，单个样本最长用时为 2 小时 1 分钟。在有、无网络搜索两种条件下，同一样本的识别结果接近，在特征类型细分情况上（比如孔特征的通盲判断）各有优劣。

测试中出现的典型错误是：GPT-6 会将部分模型的非特征面组误识别为加工特征，导致错误识别了过多特征。对于复杂模型，GPT-6 存在将两个圆孔识别为同一个特征和将圆柱面误判为圆孔等问题，见图 4。



两个圆孔被识别为同一个特征
橙色区域在空间上彼此分离



圆柱面被误识别为圆孔
橙色区域实际位于实体外侧

图 4 复杂模型中的特征实例划分与类别判断错误

总体而言，GPT-6 目前的能力是：

- 对于特征相互独立或仅存在简单嵌套关系的模型，GPT-6 能够准确识别孔、槽、型腔、圆角等全部加工特征。

- 对于包含多层嵌套、相交或截断结构的模型，GPT-6能正确识别大部分加工特征，但存在少量错误：包括（1）对于整体外形复杂的模型，GPT-6容易将非特征的表面面组误识别为特征；（2）对于结构密集、遮挡明显的模型，GPT-6对其内部特征的识别仍存在问题。

（3）工程图到 CAD 模型重建/生成

工程图到 CAD 模型重建/生成任务考察 GPT-6 能否综合多视图、剖视图和尺寸标注，重建零件的三维几何，并使分别生成的零件满足装配要求。测试包含 5 个单零件样本和 3 个多零件样本，简单、中等和复杂样本分别为 2 个、3 个和 3 个，代表样本的参考模型见图 5。

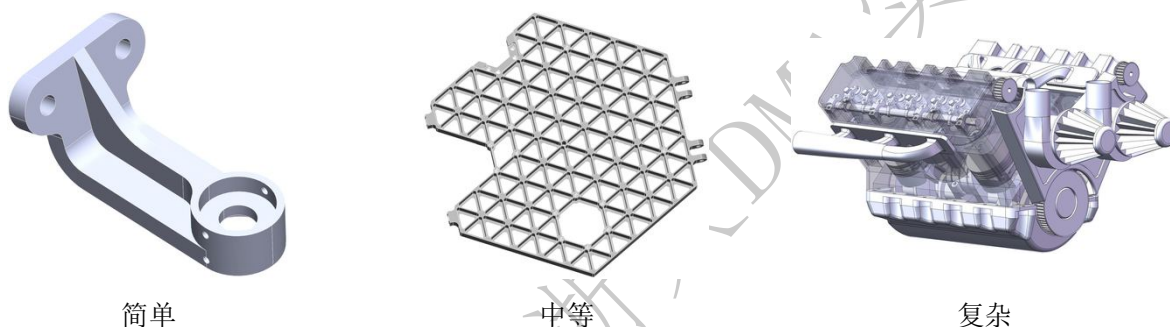


图 5 工程图到 CAD 模型重建/生成任务的不同复杂度样本

测试结果显示，实验的 8 个样本中，6 个样本的全部零件通过实体有效性检测；其余 2 个样本未通过，存在自交、非流形连接等错误。3 个多零件样本中，1 个能够完成装配，1 个存在零件干涉，1 个因结构缺失和尺寸错误无法完成装配。简单、中等和复杂样本的平均用时分别为 26 分 30 秒、43 分 40 秒和 31 分 20 秒，单个样本最长用时为 1 小时 8 分钟。

测试中出现的典型错误是：GPT-6 进行局部结构重建时会出现偏差，以及多零件配合任务存在参数偏差，使得部分配合尺寸不一致，造成零件实体重叠等问题，见图 6。

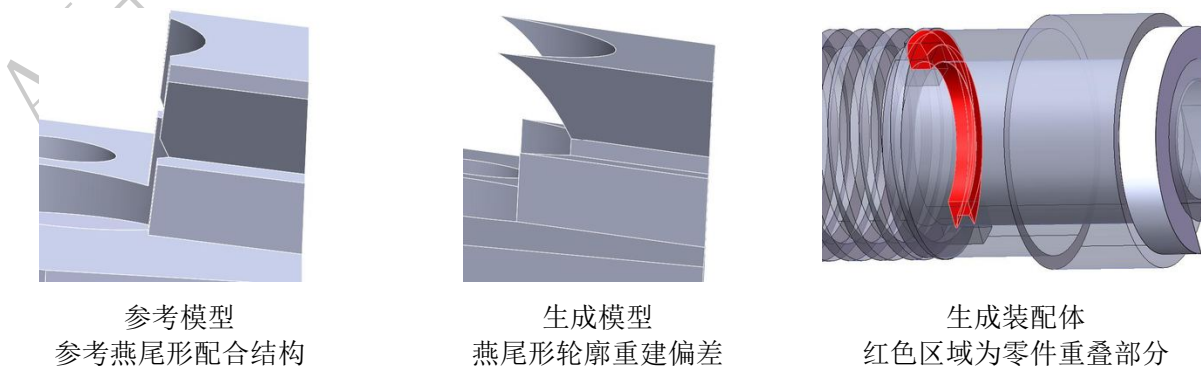


图 6 燕尾形配合结构重建偏差与装配干涉

总体而言，GPT-6 目前的能力是：

- 对于视图关系清晰、以常规孔槽和规则阵列为主的零件，GPT-6能够较好地恢复主体几何；但对于包含多层台阶、嵌套型腔及复杂薄壁过渡的零件，GPT-6仍容易误判结构层次、遗漏局部特征，有时还存在实体无效问题。
- 对于零件较少、配合关系明确的装配体，能够生成可完成装配的零件模型；但对于零件较多、配合要求复杂的装配体，GPT-6难以准确还原关键配合尺寸，容易造成干涉或无法装配。
- 对于拔模、抽壳、放样等CAD操作，GPT-6只会用简单平面去近似，在这方面与传统CAD建模仍存在差距。

（4）文本到 CAD 模型生成

文本到 CAD 模型生成任务考察 GPT-6 能否依据几何要求或功能需求，确定结构方案与尺寸参数，并生成相应的 CAD 模型。测试选取简单、中等和复杂样本各 4 个，另以 2 个复杂装配体进行压力测试，并对 4 个复杂样本开展无网络搜索测试，代表样本的参考模型见图 7。

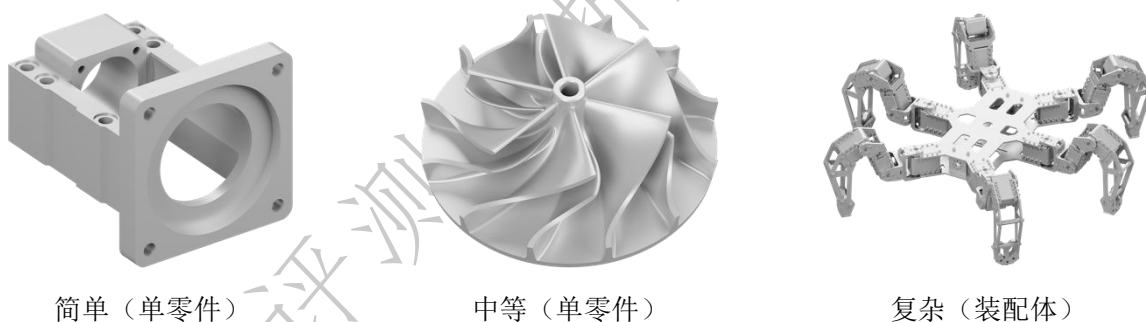
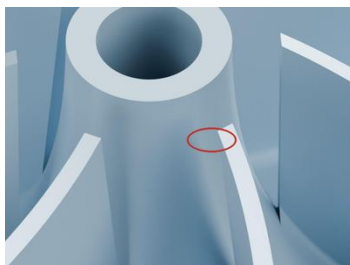


图 7 文本到 CAD 模型生成任务的不同复杂度样本

测试结果显示，12 个样本均满足全部的文本要求，11 个通过实体有效性检测。简单、中等和复杂样本平均用时分别为 42 分钟、1 小时 0 分 15 秒和 1 小时 37 分 30 秒，最长为 3 小时 17 分钟。2 个样本的压力测试中，1 个样本满足全部要求，2 个均未通过实体有效性检测；关闭网络搜索后的 4 个样本测试中，4 个均满足全部要求，2 个通过有效性检测。

测试中出现的典型错误是：GPT-6 会在本应连续连接的部位留下间隙，也会添加不承担支撑、连接或传动功能的多余构件，见图 8。



局部连接不完整
红圈处存在连接间隙



生成多余构件
橙色构件不承担实际功能

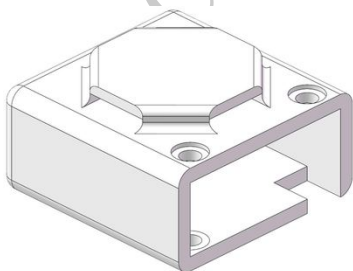
图 8 文本建模中的典型错误

总体而言，GPT-6 目前的能力是：

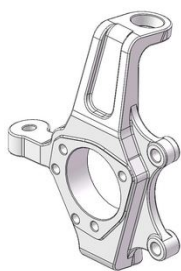
- 对于尺寸和特征位置明确的零件，或载荷、运动及安装要求清晰的机构，GPT-6能够依据文本确定结构方案和尺寸参数，生成单零件及装配体，并在迭代中补齐特征、调整部件位置（12个样本平均迭代4次）。
- 对于需要曲面连续过渡或零件协同运动的模型，GPT-6仍会在应接合处留下间隙，或使传动部件缺少固定、彼此分离，导致运动无法传递。GPT-6还会生成没有实际用途的多余构件，表明其对部件配合和构件用途的判断仍有不足，不能可靠落实文本中隐含的工程要求。
- 对于具有薄壁结构，以及内部孔腔交汇、孔壁贴近外壁的零件，GPT-6仍不能可靠处理曲面形状与相邻面的边界关系，生成结果会出现曲面自交以及非流形连接等问题。

（5）图像到 CAD 模型重建/生成

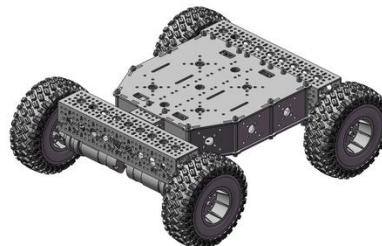
图像到 CAD 模型重建/生成任务以无尺寸标注的多视图渲染图为输入，考察 GPT-6 对三维形状、特征位置及连接关系的恢复能力。测试选取简单、中等和复杂样本各 2 个，另以 1 个复杂壳体开展压力测试，并设置不同视图组合和无网络搜索条件，代表样本见图 9。



简单



中等



复杂

图 9 图像到 CAD 模型重建/生成任务的不同复杂度样本

实验结果显示，21 次测试均生成 STEP 模型，输出中的各个实体均通过实体有效性检查。全部实验中，简单、中等和复杂样本的平均用时分别为 12.5 分钟、约 16.33 分钟和 43 分钟。采用六视图加轴测图时，6 个样本的多视图轮廓平均交并比（IoU）为 0.8790～0.9915，GPT-6 较好地恢复了主体轮廓和部件布局，但局部结构与实际尺寸仍有明显偏差。

测试中出现的典型错误是：对于包含自由曲面、薄壁及筋板交汇的零件，GPT-6 容易将复杂曲面简化为平面或分段曲面，并误判局部特征的方向和连接关系。例如，图 10 中的控制臂重建结果虽保留了总体轮廓，但侧壁的光顺曲面被分段曲面替代，根部圆角缺失，局部形成台阶。对于单实体模型，GPT-6 有时会生成多个相互穿插、体积重叠的实体，连接方式与参考零件不符。

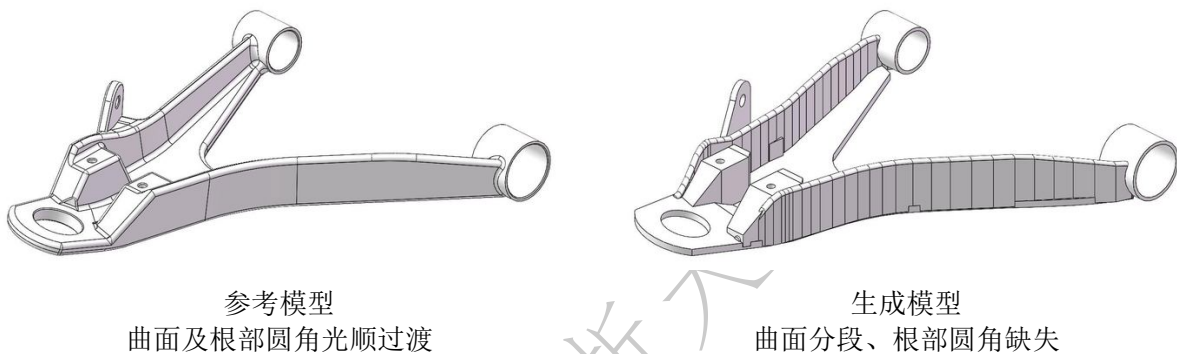


图 10 控制臂曲面及根部过渡的重建偏差

总体而言，GPT-6 目前的能力是：

- 对于轮廓清晰、主要特征充分可见的规则板件和壳体，GPT-6 能够较好恢复主体形状与孔槽布局。
- 对于自由曲面、薄壁及密集筋板交汇的结构，GPT-6 仍容易简化曲面和连接过渡，难以准确还原曲面与圆角。
- 对于复杂装配体，GPT-6 生成的模型实体间连接、孔槽及装配接口会有偏差。
- 对于无尺寸参照的建模情况，GPT-6 难以准确确定实际尺度甚至比例。

(6) B-rep 模型参数化重建

B-rep 模型参数化重建任务考察 GPT-6 能否依据已有模型的几何与拓扑信息，重建草图、建模特征及参数依赖关系，生成几何一致、可编辑的 CAD 模型。测试选取简单样本 5 个、中等样本 3 个和复杂样本 3 个，并对 3 个复杂样本进行无网络搜索测试，代表样本见图 11。

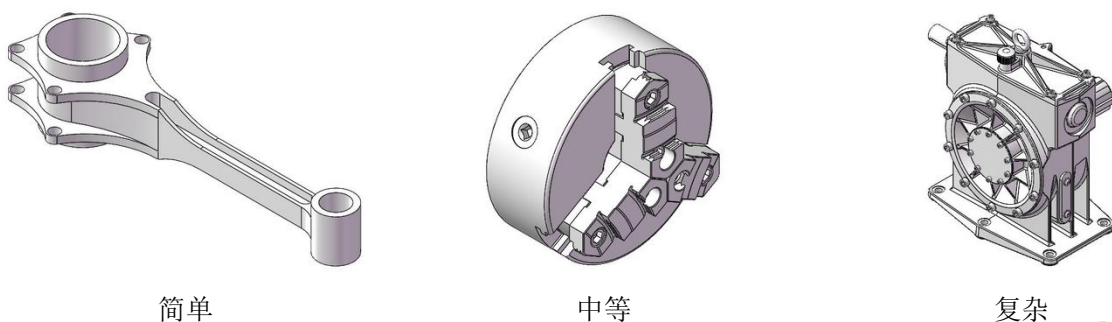


图 11 B-rep 模型参数化重建任务的不同复杂度样本

实验结果显示，实验的 11 个样本均生成了参数化模型（SolidWorks 格式），其中 6 个（占比 54.5%，5 个简单、1 个中等）通过实体有效性检查，具备参数编辑测试条件。其余 2 个中等模型和 3 个复杂模型中均存在无效实体，记为“无法编辑”，不进行参数编辑测试。对通过实体有效性检查的 6 个模型，各进行 10 次单参数修改，每次只改变一个尺寸参数。当模型参数修改生效、相关结构正确联动、修改后的模型通过实体有效性检查且能恢复原状态时，判定为通过。60 次单参数修改中，52 次通过，按样本统计，4 个样本通过全部参数编辑测试，占全部 11 个样本的 36.4%，其中 3 个简单、1 个中等。简单、中等和复杂样本的平均用时分别为 7 分 36 秒、13 分 20 秒和 50 分钟，单个样本最长用时为 1 小时 52 分钟。关闭网络搜索后的 3 个样本测试中，3 个均未通过实体有效性检查。

测试中出现的典型错误是：GPT-6 重建的模型虽然具有草图、特征和尺寸参数，外形也与参考模型接近，但仍然存在参数依赖关系不完整、可编辑性不足等问题。例如，壳体深度增加后，相关孔特征未随之正确更新，原有孔口被新增材料封闭。对于含复杂曲面的模型，GPT-6 常用多段拉伸形成的平面近似原有曲面，使光滑表面出现密集分段和台阶，造成较大的几何误差，所构建的特征结构不符合此类零件的常规建模方式，见图 12。

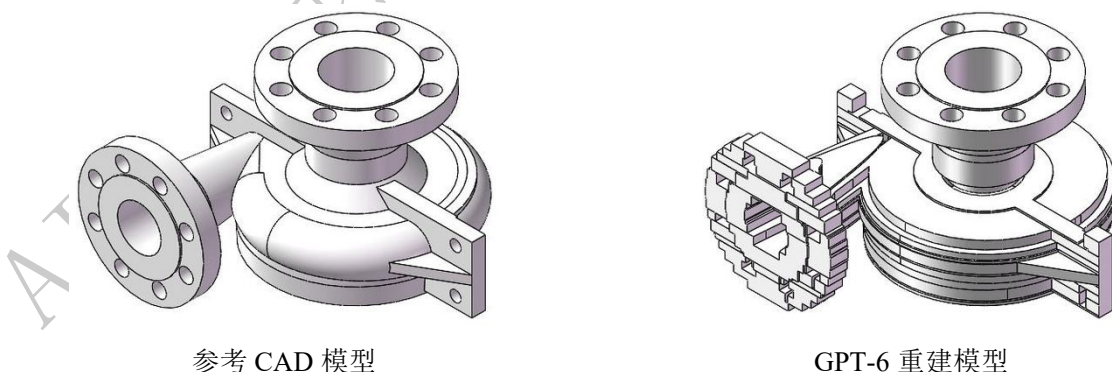


图 12 GPT-6 利用多段拉伸平面近似原有曲面

总体而言，GPT-6 目前的能力是：

- 对于以拉伸、旋转和孔槽等常规特征为主、尺寸关系较简单的模型，GPT-6 能够较准确地恢复主体几何，并支持部分尺寸修改。

- 对于孔槽与主体尺寸相互关联的模型，GPT-6建立的特征依赖关系仍可能不完整，修改主体尺寸后会出现孔口被封闭、安装孔消失等问题。
- 对于包含复杂曲面的模型，GPT-6常用多段拉伸的简单平面近似原有曲面，原本光滑的表面出现分段和台阶，重建误差较大。

AI+CAD评测，浙大DMA实验室

目录

摘要	1
1 测试目标与总体设计	13
1.1 测试目标	13
1.2 测试任务范围	13
1.3 测试环境配置	14
2 零件分类	15
2.1 任务目标与评价标准	15
2.2 测试用例设计	15
2.3 测试结果展示	19
2.4 GPT-6 推理过程分析	29
2.5 能力边界	30
3 特征识别	31
3.1 任务目标与评价标准	31
3.2 测试用例设计	31
3.3 测试结果展示	34
3.4 GPT-6 推理过程分析	42
3.5 能力边界	44
4 工程图到 CAD 模型重建/生成	46
4.1 任务目标与评价标准	46
4.2 测试用例设计	46
4.3 测试结果展示	50
4.4 GPT-6 推理过程分析	68
4.5 能力边界	70
5 文本到 CAD 模型生成	72
5.1 任务目标与评价标准	72

5.2 测试用例设计	72
5.3 测试结果展示	79
5.4 GPT-6 推理过程分析	104
5.5 能力边界	106
6 图像到 CAD 模型重建/生成	108
6.1 任务目标与评价标准	108
6.2 测试用例设计	109
6.3 测试结果展示	112
6.4 GPT-6 推理过程分析	148
6.5 能力边界	150
7 B-rep 模型参数化重建	151
7.1 任务目标与评价标准	151
7.2 测试用例设计	151
7.3 测试结果展示	154
7.4 GPT-6 推理过程分析	171
7.5 能力边界	173
8 综合分析	174

1 测试目标与总体设计

1.1 测试目标

本报告系统评估OpenAI GPT-6在CAD模型理解与生成任务上的能力，主要关注其不同任务复杂度下的表现、推理过程和能力边界。报告可用于辅助判断GPT-6在哪些CAD任务上已接近实用水平、哪些仍存在能力瓶颈，以及哪些研究方向最值得后续投入。

- 模型理解：评估GPT-6对CAD模型的几何、拓扑和工程语义的理解能力。
- 模型生成：评估GPT-6从工程图、文本和渲染图像生成CAD模型的能力，以及从CAD模型中恢复参数化建模历史的能力。
- 分析GPT-6完成不同任务时采用的可能推理过程，以及在不同任务上的能力边界。
- 基于测试结果形成整体能力判断，启发未来智能CAD研究方向。

1.2 测试任务范围

本次测试具体包括6类任务。每个任务均沿用“测试用例设计 → 结果展示 → GPT-6推理过程分析 → 能力边界”的统一框架。不同任务内部设计对应的复杂度标准选取测试数据、输入Prompt和评价指标。具体任务内容如表1-1所示：

表1-1 测试任务总览

任务	输入	期望输出	主要观察能力
零件分类	B-rep模型（STEP文件）	零件类别、判断依据、典型功能用途	整体几何/拓扑理解与语义识别
特征识别	B-rep模型（STEP文件）	加工特征和局部面对应关系	局部几何/拓扑理解、特征语义
工程图到CAD	二维工程图及相关装配资料	单零件CAD模型或装配体全部零件的CAD模型	工程图理解、几何与尺寸推理、装配关系理解和CAD模型重建能力
文本到CAD模型生成	几何要求或功能需求描述文本	满足文本中几何要求或功能需求的CAD模型	几何要求理解与实现和功能需求到模型特征的转化。
图像到CAD模型重建/生成	产品多视角渲染图	与可见外形及主要结构一致的CAD模型	多视图理解能力、模型特征与尺寸参数/比例推理、局部特征重建及不确定性处理

任务	输入	期望输出	主要观察能力
B-rep模型参数化重建	B-rep模型（STEP文件）	与输入几何一致、具有建模历史的可编辑参数化CAD模型	几何与拓扑理解、建模特征与草图推断、建模顺序规划及参数化编辑能力

1.3 测试环境配置

所有测试任务采用统一GPT-6环境配置，见表1-2。

表1-2 GPT-6环境配置表

配置项	配置内容
模型	GPT-6 Astra
推理强度	MAX
使用模式	ChatGPT客户端，Codex模式远程连接Linux服务器
会话策略	每个测试任务使用独立会话，避免其他对话泄露标签或解法
记忆	关闭
文件输入能力	允许读取STEP文件、图像及任务所需必要附件
代码执行	允许GPT-6自主创建环境
网络搜索	允许网络搜索，也进行了关闭搜索的测试实验
CAD/几何工具	未预设外部CAD/几何工具；GPT-6可能自主调用可用工具，具体流程见推测分析
迭代限制	只进行单轮对话，允许单轮内自主迭代

2 零件分类

2.1 任务目标与评价标准

零件分类任务主要测试GPT-6能否依据零件B-rep模型（STEP文件）所包含的几何与拓扑信息，在预先给定的零件类别集合中判断其所属类别。该设置与传统零件分类任务保持一致，如此可获得可比较的分类结果。

在此基础上，进一步通过开放式工程语境相关的Prompt，扩展标准分类任务，考察GPT-6能否脱离固定标签集合，对零件名称、结构依据、功能用途及工程语义进行更深入的理解与推断。具体任务说明见表2-1。

表2-1 零件分类任务说明

测试要素	说明
测试问题	GPT-6能否仅依据B-rep模型信息，对机械零件进行正确分类，并在不依赖固定标签集合的情况下进一步形成对零件类别、结构依据及典型功能用途的理解？
核心输入	机械零件的B-rep模型。
核心输出	最可能的零件类别/名称、主要判断依据、典型功能用途及必要的不确定性说明。
评价标准	标准测试任务以给定标签集合下的分类正确性为主要指标；扩展测试任务则评价零件类别/名称判断的合理性、几何与结构依据是否准确、功能用途推断是否合理，以及对不确定性的识别是否恰当。

2.2 测试用例设计

2.2.1 测试样本选取

零件分类任务的测试样本来自TMCAD数据集，均为真实机械零件模型且具有明确零件类别或名称。选取的样本尽可能覆盖不同类别及不同识别难度，避免仅选择外形高度典型或者仅凭单一局部结构即可判断类别的零件。

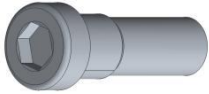
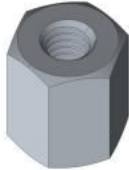
样本难度主要根据零件整体形状的典型程度、类别判别所需的几何信息范围、与其他零件类别的相似性，以及零件功能语义能否由几何/拓扑结构直接反映，具体如表2-2所示。

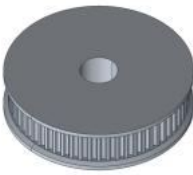
表2-2 零件分类任务样本复杂度分级表

样本复杂度	选取标准	典型表现
简单	零件整体形态和关键功能结构具有较强类别典型性，仅依据整体形状或少量关键结构即可可靠地判断类别。	类别特征明显，与相近类别差异较大。
中等	需要综合多个局部结构及其空间关系完成分类，单一局部特征不足以确定类别，并可能与相近类别具有几何相似性。	需结合主体结构、安装结构、孔系、支撑或连接结构等进行综合判断。
复杂	类别特征不直接暴露，或与其他类别存在较强的形态/结构相似性；仅凭局部结构容易产生误判。	复杂工业零件、非标准件、多功能集成零件，或存在合理类别歧义的零件。

基于上述难度分级，总共选取简单、中等和复杂三个级别的样本各4个。具体信息见表2-3。

表2-3 零件分类任务样本清单

样本ID	CAD模型渲染图	零件类别	复杂度	复杂度归类说明
PC-01		轴承	简单	单一环形主体及内侧连续滚道呈现轴承套圈特征。
PC-02		螺钉/螺栓	简单	圆柱杆身、加大圆柱头及内六角驱动孔构成典型螺栓形态。
PC-03		支架	简单	三个相互垂直的面均设有安装孔，转角支撑与固定关系直观。
PC-04		螺母	简单	外六角轮廓与螺纹通孔清晰可见。

样本ID	CAD模型渲染图	零件类别	复杂度	复杂度归类说明
PC-05		支架	中等	两端支撑结构、中央连接体及安装孔共同形成支撑布局，多组安装孔增加了类别判断的复杂度。
PC-06		法兰	中等	中央贯通孔与六条径向安装臂、臂端孔系组合成连接界面；仅看辐射臂易与支架混淆。
PC-07		齿轮	中等	轮缘斜齿连续扭转，局部外观近似叶片。需结合周向齿列与中心轴孔判断齿轮类别。
PC-08		带轮	中等	外周密集带齿与两侧挡边并存，单看齿圈易与齿轮混淆；需综合挡边、轴孔及传动轮形态进行判断。
PC-09		轴承	复杂	内外圈与多件环列滚动体组成嵌套结构，滚动部件部分被圈体遮挡，仅看外轮廓易误认为环套件。
PC-10		连接器	复杂	整体呈同轴圆筒状，外形与其他轴类零件相似，但内部连接结构复杂
PC-11		轴	复杂	多个轴段相互错位，整体形状与常见直轴差异较大。

样本ID	CAD模型渲染图	零件类别	复杂度	复杂度归类说明
PC-12		弹簧	复杂	两端卷圈由长折弯臂连接，缺少常见密绕螺旋外形。

2.2.2 测试条件设计

(1) 主实验设计

主实验基于第1.3节的测试环境配置，使用统一的Prompt对全部测试样本进行测试。沿用通常的零件分类任务定义，该Prompt提供预先定义的零件类别集合，要求GPT-6根据输入的B-rep模型判断零件所属类别。在具体测试中，所有样本使用相同的类别集合，仅改变测试样本，以评价GPT-6在不同样本复杂度下的零件分类能力。具体Prompt为：“在bearing bracket coupling flange gear nut pulley shaft screw_bolt spring这些类别中，给定的step模型属于哪类零件”。

(2) 扩展实验设计

在主实验基础上，我们选取代表性复杂样本进行扩展测试，主要考察Prompt形式和网络搜索权限对零件识别结果的影响。扩展实验设置两个变量：一是将原有的标准分类Prompt替换为开放理解Prompt，取消候选类别限制，并要求模型解释判断依据及典型功能用途；二是关闭网络搜索权限，观察在无法获取外部网络信息时GPT-6的表现。开放理解Prompt的内容为：“请判断这个CAD模型最可能是什么零件。请说明你做出这一判断所依据的几何特征，并说明这种零件通常具有哪些功能或用途。如果无法唯一确定，请给出你认为最可能的类别，并说明存在的不确定性”。

以上两个变量组合形成3种扩展测试条件，如表2-4所示。扩展实验仅改变Prompt和网络搜索权限，其余配置与主实验保持一致。这些设置使用独立会话进行测试，避免不同实验条件之间的信息干扰。

表2-4 零件分类任务扩展实验条件

设置	Prompt	网络搜索	测试目的
S1开放理解	P2开放理解	开启	考察无固定标签约束下的自主识别与工程语义理解。
S2关闭搜索	P1标准分类	关闭	考察标准分类结果对网络搜索的依赖情况。

设置	Prompt	网络搜索	测试目的
S3开放理解+关闭搜索	P2开放理解	关闭	考察无固定标签约束、无网络搜索条件下的零件识别与工程语义理解能力。

2.3 测试结果展示

本节按照主实验与扩展实验分别展示测试结果。主实验重点评价GPT-6在给定类别集合下的零件分类正确性，并分析不同样本复杂度下的表现。扩展实验选取代表性复杂样本，通过改变Prompt和网络搜索权限，进一步考察模型的自主识别、结构依据判断及工程语义理解能力。两类实验结果分别展示与评价。

2.3.1 主实验结果

本节首先汇总整体分类结果，随后逐样本展示零件模型、GPT-6预测类别及预测原因，重点观察分类依据。

(1) 总体概览

主实验分类全部成功，如表2-5所示。各测试样本的用时见表2-6。

表2-5 零件分类任务结果汇总表

样本ID	复杂度	真实类别	预测类别	分类结果
PC-01	简单	轴承	轴承	正确
PC-02	简单	螺钉/螺栓	螺钉/螺栓	正确
PC-03	简单	支架	支架	正确
PC-04	简单	螺母	螺母	正确
PC-05	中等	支架	支架	正确
PC-06	中等	法兰	法兰	正确
PC-07	中等	齿轮	齿轮	正确
PC-08	中等	带轮	带轮	正确
PC-09	复杂	轴承	轴承	正确
PC-10	复杂	联接器	联接器	正确
PC-11	复杂	轴	轴	正确
PC-12	复杂	弹簧	弹簧	正确

表2-6 零件分类主实验各样本测试用时

样本ID	PC-01	PC-02	PC-03	PC-04	PC-05	PC-06	PC-07	PC-08	PC-09	PC-10	PC-11	PC-12
用时	3分钟	2分钟	2分钟	3分钟	2分钟	2分钟	2分钟	2分钟	2分钟	4分钟	2分钟	2分钟

(2) 用例分析

PC-01（轴承 简单）



GPT-6页面输出

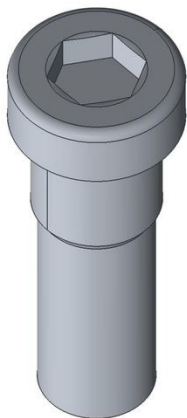
本轮完成easy_0：分类为bearing（轴承），具体为轴承外圈，依据是内侧环形凹滚道及两端台阶。

classification.json

```
case easy_0
category bearing（轴承）
component 轴承外圈
confidence high
reason 模型为同轴环形实体，外径6 mm、轴向宽2.5 mm；内侧具有圆弧截面的环形凹滚道及两端台阶，符合滚珠轴承外圈的几何特征。
validation STEP读取成功；有效模型：1个实体、16个面
```

结果分析：GPT-6正确将模型归为轴承。它进一步判断该模型为轴承外圈，并以内侧环形凹槽和两端台阶为依据；这一判断与模型结构一致。

PC-02（螺钉/螺栓 简单）



GPT-6页面输出

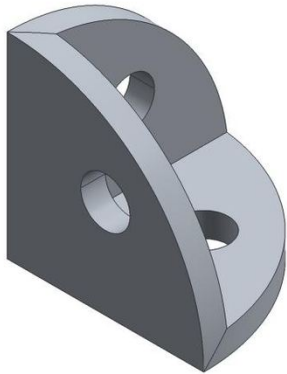
本轮完成easy_1：分类为screw_bolt（螺钉/螺栓），依据是长圆柱杆、扩大头部及内六角驱动凹槽。

classification.json

```
case easy_1
category screw_bolt（螺钉/螺栓）
component 内六角圆柱头螺钉/螺栓
confidence high
reason 模型具有长圆柱杆和扩大的圆柱头，头端六个内侧平面构成内六角驱动凹槽，符合有头紧固件特征。杆部主要圆柱直径30 mm，头部直径45 mm，总轴向长度117.5 mm。模型未显式建出螺旋牙型，分类依据为整体形态及内六角驱动结构。
validation STEP读取成功；有效模型：1个实体、22个面
```

结果分析：GPT-6正确将模型归为螺钉/螺栓，并依据杆身、头部和内六角凹槽判断。模型没有直接画出螺纹，但这些外形特征仍足以支持分类。

PC-03（支架 简单）



GPT-6页面输出

本轮完成easy_2：分类为bracket（支架），具体为三向直角连接支架，依据是三片正交板及各板上的安装孔。

classification.json

case easy_2

category bracket（支架）

component 三向直角连接支架

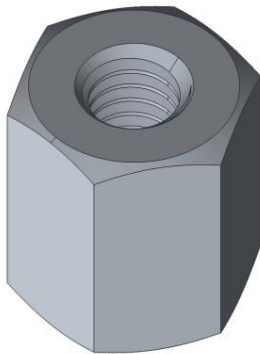
confidence high

reason 模型由三片互相垂直、厚5 mm的板组成，每片板上各有一个直径10 mm的贯穿安装孔，三个孔轴互相正交；整体包围尺寸40×40×40 mm。该三向安装连接结构符合角支架特征。

validation STEP读取成功；有效模型；1个实体、12个面

结果分析：GPT-6正确将模型归为支架。它注意到三个方向的板和安装孔，能利用清晰的整体结构进行判断。

PC-04（螺母 简单）



GPT-6页面输出

本轮完成easy_3：分类为nut（螺母），依据是六角柱外形和贯穿内螺纹孔。

classification.json

case easy_3

category nut（螺母）

component 六角内螺纹隔离柱（长螺母形）

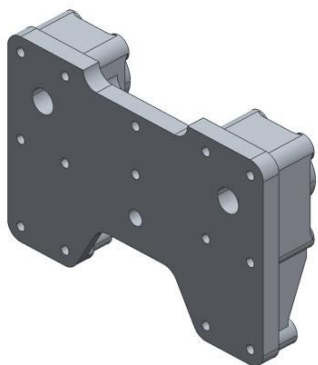
confidence high

reason 模型为对边5 mm、轴向长6 mm的六角柱，中心具有贯穿内螺纹孔，两端有倒角。六角扳拧外形与内螺纹连接结构符合螺母类特征，在给定类别中归入nut。

validation STEP读取成功；有效模型；1个实体、44个面

结果分析：GPT-6正确将模型归为螺母。六角外形和中心螺纹孔是直接的判断依据，并给出了更具体的名称：六角内螺纹隔离柱。

PC-05（支架 中等）



GPT-6页面输出

本轮完成medium_0：分类为bracket（支架），依据是带加强结构的非回转板体及多处安装孔、凸台。

classification.json

case medium_0

category bracket（支架）

component 多孔承载安装支架

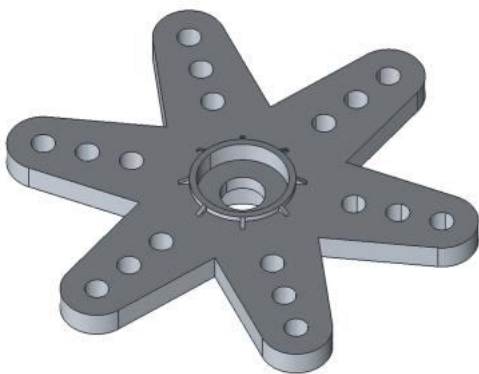
confidence high

reason 模型为非回转的板体与加强结构，具有多个安装凸台及延伸连接部位。两个直径16 mm的平行主孔相距142 mm，周边和中央另分布14个直径6 mm的孔，并有局部沉孔。整体几何符合用于承载、定位和安装连接的支架。

validation STEP读取成功；有效模型；1个实体、179个面

结果分析：GPT-6正确将模型归为支架。它没有受到大量孔特征的影响，而是同时参考板体、加强结构及多处安装凸台进行了正确判断。

PC-06（法兰 中等）



GPT-6页面输出

本轮完成medium_1：分类为flange（法兰）。模型具体为六臂舵盘，依据中央轮毂及径向安装孔，在给定类别中归入法兰。

classification.json

case medium_1

category flange（法兰）

component 六臂舵盘（异形轴端连接盘）

confidence medium_high

reason 模型中央具有带阶梯孔的短圆柱轮毂，周围六个等角分布的径向薄臂上各有三个轴向安装孔，共18孔，臂厚1.8 mm、安装孔直径1.6 mm。该结构通过中央轮毂与轴端连接、通过径向孔连接外部构件，整体最接近异形连接法兰。

validation STEP读取成功；有效模型；1个实体、72个面

结果分析：GPT-6正确将模型归为法兰。它在判断时综合考虑了中央连接部分和向外分布的安装孔，并进一步指出模型实际上是六臂舵盘。

PC-07（齿轮 中等）



GPT-6页面输出

本轮完成medium_2：分类为gear（齿轮），具体为20齿斜齿圆柱齿轮。

classification.json

case medium_2

category gear（齿轮）

component 20齿斜齿圆柱齿轮

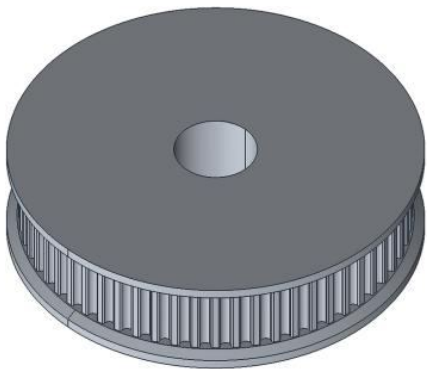
confidence high

reason模型外周具有20组等角重复的齿形，齿面沿轴向明显螺旋倾斜，形成斜齿圆柱齿轮。齿根圆半径17.5 mm，外径约44 mm，轴向厚20 mm，中央贯穿轴孔直径12 mm；几何拓扑及模型视图均支持齿轮分类。

validation STEP读取成功；有效模型；1个实体、143个面

结果分析：GPT-6正确将模型归为齿轮。它能发现外周连续排列的齿形，并进一步指出斜齿结构。

PC-08（带轮 中等）



GPT-6页面输出

本轮完成medium_3：分类为pulley（带轮），具体为60齿同步带轮，依据是细齿带面及两侧挡边。

classification.json

case medium_3

category pulley（带轮）

component 60齿同步带轮

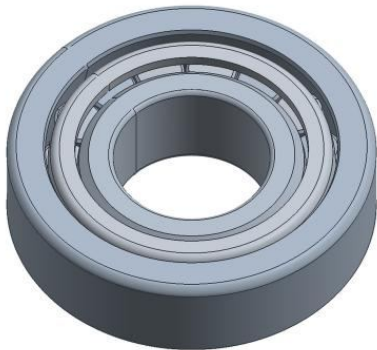
confidence high

reason模型外周具有60个等角重复的细齿槽，齿带宽7 mm，两侧有直径40 mm的挡边；中央贯穿轴孔直径8 mm，轮毂直径20 mm，并有两个互相正交的直径3 mm径向紧定孔。细齿带面与两侧挡边构成典型同步带轮结构。

validation STEP读取成功；有效模型；1个实体、326个面

结果分析：GPT-6正确将模型归为带轮，并指出是同步带轮。虽然轮缘上的齿容易让人联想到齿轮，但它结合细齿及两侧挡边给出了更具体的判断。

PC-09（轴承 复杂）



GPT-6页面输出

本轮完成hard_0: 分类为bearing（轴承），依据是内外圈、保持架及周向均布的15组滚子。

classification.json

case hard_0

category bearing（轴承）

component 滚子轴承组件

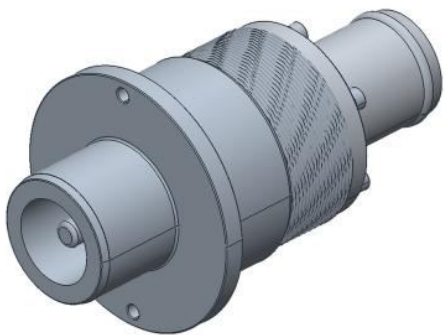
confidence high

reason 模型由33个实体组成，包括同轴内圈、外圈、带周向窗口的环形保持架，以及15组均布的滚子和销。内圈轴孔直径17 mm、外圈外径40 mm；滚动体位于内外圈之间并由保持架定位，明确符合滚动轴承结构。

validation STEP读取成功；有效模型；33个实体、298个面

结果分析：GPT-6正确将模型归为轴承。文件记录显示模型含内外圈、保持架和滚子等多个部分；回答能结合这些部件判断，但从渲染外观上看，轴承类别本身也比较明显。

PC-10（管路快速接头 复杂）



GPT-6页面输出

本轮完成hard_1: 分类为coupling（连接器），具体为带弹簧阀芯的管路快速接头公端。

classification.json

case hard_1

category coupling（连接器）

component 带弹簧阀芯的管路快速接头（公端）

confidence high

reason 模型为11个实体组成的同轴接头装配体，包含筒状主体、带轴向流道的连接端、六颗周向均布紧固件，以及内部阀芯、导向件和弹簧。阶梯插接端与内部通道、阀机构共同构成管路快速联接结构，应按整体归入coupling。模型自身原始名称Quick coupler (male)也与几何判断一致。

validation STEP读取成功；有效模型；11个实体、266个面

结果分析：GPT-6将模型归为coupling，并将coupling译为连接器，不影响分类结果的正确性。它也识别出具体零件是管路快速接头，而非通常连接两根转轴的联轴器。其内部弹簧和阀芯为进一步辨认提供了关键依据。

PC-11（轴 复杂）



GPT-6页面输出

本轮完成hard_2：分类为shaft（轴），具体为 四拐曲轴，依据是同轴主轴颈、偏心连杆轴颈及曲柄结构。

classification.json

```
case hard_2
category shaft（轴）
component 四拐曲轴
confidence high
reason 模型具有5段同轴主轴颈（半径36 mm）及4个偏心连杆轴颈（半径27 mm、偏心量44 mm、各长46 mm），各偏心轴颈两侧通过曲柄和平衡重与主轴连接，另有细长轴端。该多拐偏心传动结构明确属于曲轴，归入shaft。
validation STEP读取成功；有效模型；1个实体、61个面
```

结果分析：GPT-6正确将模型归为轴，并进一步识别出四拐曲轴。它没有因为多个轴段错位就忽略整体连接关系，而是结合偏心轴段及曲柄结构进行判断。

PC-12（弹簧 复杂）



GPT-6页面输出

本轮完成hard_3：分类为spring（弹簧），具体为双扭转弹簧。

classification.json

```
case hard_3
category spring（弹簧）
component 双扭转弹簧
confidence high
reason 模型由直径2 mm的连续圆线构成，左右各有盘绕段及长力臂，两侧通过跨接弯曲段连接，末端为圆截面。左右对称的双线圈及力臂结构明确符合双扭转弹簧；几何解析与模型视图一致。
validation STEP读取成功；有效模型；1个实体、15个面
```

结果分析：GPT-6正确将模型归为弹簧，并给出双扭转弹簧的具体名称。尽管模型外形不同于常见的螺旋弹簧，GPT-6依据两侧盘绕部分和伸出的力臂进行了正确判断。

2.3.2 扩展实验结果

本节依照第2.2.2节的S1、S2和S3设置分别展开：每项实验均以相同样本在主实验中的输出为基线，从复杂级样本中选取具有解释价值的代表性用例进行展示。各测试样本的用时见表2-7。

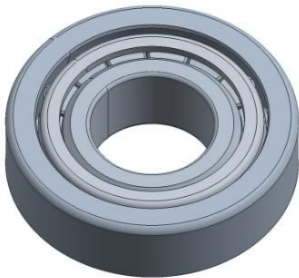
表2-7 零件分类扩展实验各样本测试用时

样本ID	S1_PC-09	S1_PC-10	S2_PC-10	S2_PC-12	S3_PC-09	S3_PC-10
用时	11分钟	17分钟	1分钟	3分钟	7分钟	7分钟

(1) S1开放理解

本设置保持网络搜索权限不变，取消候选类别列表，并要求说明几何依据、用途及不确定性。选取PC-09观察细分类中的几何核验，选取PC-10观察固定标签之外的具体产品语义。

PC-09（轴承）

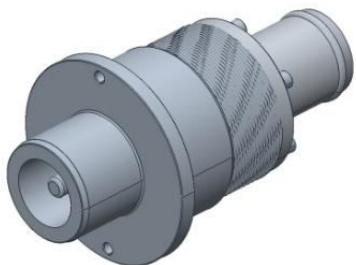


主实验输出

分类：bearing（轴承）；具体为滚子轴承组件。
依据：模型由33个实体组成，包括同轴内圈、外圈、带周向窗口的环形保持架，以及15组均布的滚子和销。内圈轴孔直径17 mm、外圈外径40 mm；滚动体位于内外圈之间并由保持架定位，明确符合滚动轴承结构。
S1输出
判断：滚子轴承装配体，最可能是单列圆锥滚子轴承的简化或非标准模型。
依据：内外圈具有锥形滚道，15个主要滚子单列倾斜排列；但滚子主工作面没有锥度，主体曲面均为半径2.75 mm的圆柱。
用途及保留：滚子轴承用于支承旋转轴，并通过滚动接触传递支承载荷；对“滚子轴承装配体”的判断置信度高，对“标准单列圆锥滚子轴承”的细分类置信度为中等。

对比分析：主实验输出bearing和“滚子轴承组件”，S1仍确认滚子轴承大类，但将细分类推进到“单列圆锥滚子轴承的简化或非标准模型”，同时通过圆柱滚子的解析几何指出反例，并明确将标准细分类置信度限定为中等。S1的增量主要体现在具体子类型、几何核验与不确定性解释。

PC-10（管路快速接头）



主实验输出

分类: coupling（连接器）；具体为带弹簧阀芯的管路快速接头（公端）。

依据: 模型为11个实体组成的同轴接头装配体，包含筒状主体、带轴向流道的连接端、六颗周向均布紧固件，以及内部阀芯、导向件和弹簧。阶梯插接端与内部通道、阀机构共同构成管路快速联接结构，应按整体归入coupling。

S1输出

判断: 带弹簧自闭阀的水路接头总成；进一步判断为水路公端快速接头。

依据: 11个有效实体组成的同轴流道壳体，包含阀杆、锥面阀瓣、导向件和复位弹簧。

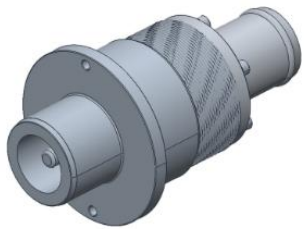
用途及保留: 通常用于需要拆接的水路或冷却水管路；缺少配合接头，不能确定具体锁紧、开启方式或接口标准，也不能从静态形状确认密封效果。

对比分析: 主实验在给定标签中归入coupling，且已识别“带弹簧阀芯的管路快速接头（公端）”；S1同样将整个装配判断为水路公端快速接头，并增加水路用途与缺少配合件时不能确认锁紧、开启方式的说明。

(2) S2关闭搜索

本设置保留与主实验相同的标准分类Prompt和候选集合，仅关闭网络搜索。选取PC-10考察装配体在离线条件下的类别判断，选取PC-12检验非典型弹簧外形的识别稳定性。

PC-10（管路快速接头）



主实验输出

分类: coupling（连接器）；具体为带弹簧阀芯的管路快速接头（公端）。

依据: 模型为11个实体组成的同轴接头装配体，包含筒状主体、带轴向流道的连接端、六颗周向均布紧固件，以及内部阀芯、导向件和弹簧。阶梯插接端与内部通道、阀机构共同构成管路快速联接结构，应按整体归入coupling。

S2输出

分类: coupling；中文名称为连接器（快速接头）。

依据: 模型由沿同一轴线配合的阶梯筒状主体、内部构件和周向分布的紧固件组成，符合接头组件结构。

对比分析: 主实验与S2均将模型归为coupling，并识别为管路快速接头。说明PC-10成功分类并不需要依赖网络搜索功能。但相比之下，主实验的判断依据描述更加具体。

PC-12（弹簧）



主实验输出

分类：spring（弹簧）；具体为双扭转弹簧。

依据：模型由直径2 mm的连续圆线构成，左右各有盘绕段及长力臂，两侧通过跨接弯曲段连接，末端为圆截面。左右对称的双线圈及力臂结构明确符合双扭转弹簧；几何解析与模型视图一致。

S2输出

分类：spring；中文名称为弹簧（双扭转弹簧）。

依据：模型由连续圆截面线材成形，具有两侧对称的扭转线圈及伸出臂，符合双扭转弹簧的典型结构，应归入spring类。

对比分析：主实验给出spring及“双扭转弹簧”，S2也给出相同大类和具体类型，并以连续圆截面线材、两侧扭转线圈及伸出臂作为判断依据。说明PC-12成功分类并不需要依赖网络搜索功能。但相比之下，无网络条件的判断依据描述更为简单，比如不再包含尺寸信息。

(3) S3开放理解+关闭搜索

本设置同时取消固定标签和网络搜索权限。选取PC-09观察无网络条件下对轴承细分类的判断与保留，选取PC-10观察无网络条件下对装配体的工程语义识别程度。

PC-09（轴承）



主实验输出

分类：bearing（轴承）；具体为滚子轴承组件。

依据：模型由33个实体组成，包括同轴内圈、外圈、带周向窗口的环形保持架，以及15组均布的滚子和销。内圈轴孔直径17 mm、外圈外径40 mm；滚动体位于内外圈之间并由保持架定位，明确符合滚动轴承结构。

S3输出

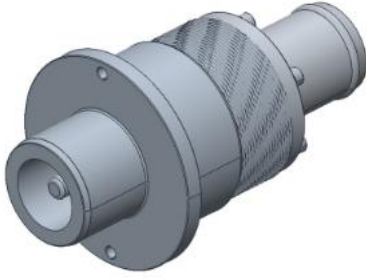
判断：最可能是单列圆锥滚子轴承总成；能够较高置信度确认的类别是“带保持架的单列滚子轴承”。

依据：内圈、外圈、保持架与一圈滚子构成轴承装配；文件含33个有效实体，其中包括15个滚子主体。滚子主体为等径圆柱面，而非圆锥面。

用途及保留：用于支承旋转轴、传递载荷；按最可能的圆锥滚子轴承理解，可承受径向和单向轴向载荷。对“滚子轴承总成”的判断置信度高，对“标准单列圆锥滚子轴承”的严格细分类置信度中等。

对比分析：主实验判断bearing，可能受到预定义类别的影响直接判断具体分类为“滚子轴承组件”；GPT-6在无网络条件下仍能直接给出细粒度的分类信息：首先确认轴承装配大类，然后提出单列圆锥滚子轴承作为细分类别候选，最后基于滚子主面为圆柱面给出中等置信度。

PC-10（管路快速接头）



主实验输出

分类：coupling（联接器）；具体为带弹簧阀芯的管路快速接头（公端）。
依据：模型为11个实体组成的同轴接头装配体，包含筒状主体、带轴向流道的连接端、六颗周向均布紧固件，以及内部阀芯、导向件和弹簧。阶梯插接端与内部通道、阀机构共同构成管路快速联接结构，应按整体归入coupling。

S3输出

判断：最可能是弹簧加载的轴向提升式止回阀（单向阀）总成，采用带导向阀杆的锥面阀瓣。

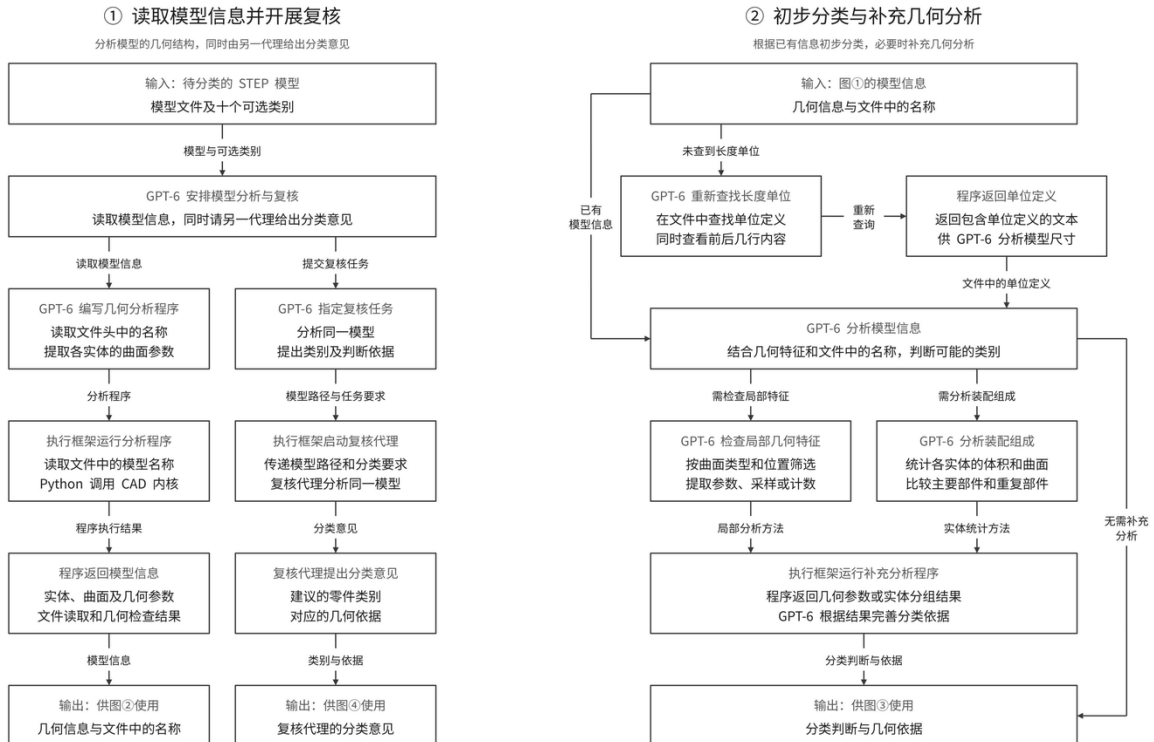
依据：模型含11个实体；两个同轴接口、锥面阀瓣、导向支架和复位弹簧支持弹簧阀机构的判断。

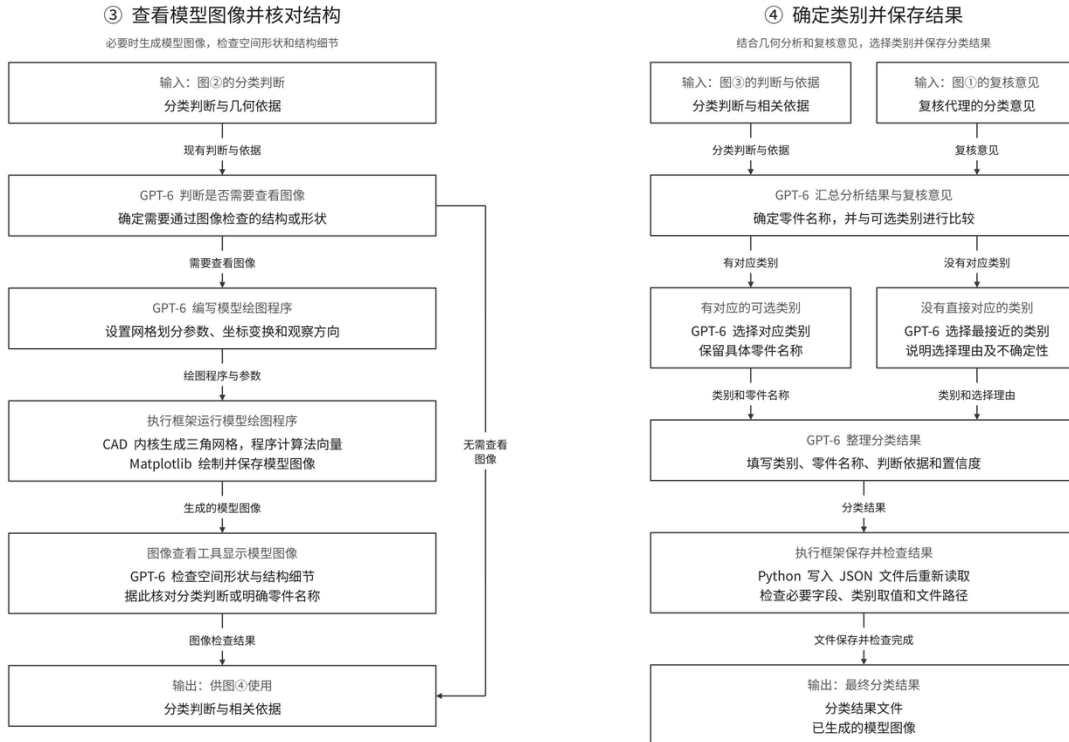
用途及保留：这类止回阀用于允许流体单向通过，并在压差不足或出现反向流动趋势时自动关闭；也可能是自封式流体接头中的止回阀机构，缺少配套接头和装配环境，无法唯一确定完整产品类型。

对比分析：主实验成功归入coupling并具体识别为管路快速接头；而S3称其为“轴向提升式止回阀（单向阀）总成”。无网络条件下，GPT-6也发现了弹簧、导向阀杆和阀瓣等内部机制，但最后只识别到了内部阀结构的信息并作为最终输出，未能对整体进行零件类别判断与功能用途推断。

2.4 GPT-6推理过程分析

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对零件分类任务可能采用如下推理流程：





2.5 能力边界

在零件分类任务上，GPT-6已能够借助CAD内核查询和必要的模型视图核查，综合零件的整体形态、局部几何特征及装配组成进行分类。在给定类别的主实验中，GPT-6对全部样本均完成了正确判断。其识别能力也不局限于典型外形；对于外形不典型或与其他类别较为相近的零件，GPT-6仍能够结合多个结构特征进行区分，并给出较为具体的名称和功能解释。关闭网络搜索后，固定类别条件下的测试结果同样保持正确。

当前阶段，GPT-6主要能力边界在于进一步确定零件的细分类型。对于装配体，当缺少配合件和装配环境等信息时，GPT-6有时会依据局部组件的结构和功能直接判断整个零件的类别，从而产生误判；当整体形态与局部几何特征之间存在不完全一致的情况时，GPT-6也往往只能给出可能的子类型，难以可靠确认其严格归属。

总体来看，GPT-6对零件的类别级识别已经较为成功，但更细粒度的类型识别仍不够可靠，主要瓶颈在于如何将几何信息进一步准确对应到零件的功能语义和具体类型。

3 特征识别

3.1 任务目标与评价标准

特征识别任务主要测试GPT-6能否仅依据B-rep模型中的几何与拓扑信息，自主识别加工特征，并按照“从高层级类别到具体子类型”进行层级化判断。测试不预先给定特征名称集合。在此基础上，进一步考察GPT-6对复杂特征干涉关系的理解，包括相交、嵌套、截断、共享边界等情况，以及能否将识别结果正确对应到具体几何区域。具体任务说明见表3-1。

表3-1 特征识别任务说明

测试要素	说明
测试问题	GPT-6能否仅依据B-rep模型的几何与拓扑信息，准确识别加工特征，并在不依赖固定标签列表的情况下完成从高层级类别到具体子类型的判断，同时正确处理复杂特征之间的相交、嵌套等关系。
核心输入	包含一个或多个加工特征的B-rep模型。
核心输出	识别出的特征结果，包括高层级类别、具体子类型、主要判断依据、特征间关系及必要的不确定性说明。
评价标准	高层级类别与子类型识别正确性和不确定性处理等方面评价；复杂样本重点关注特征相交、嵌套等情况对应区域的漏检和误判。

3.2 测试用例设计

3.2.1 测试样本选取

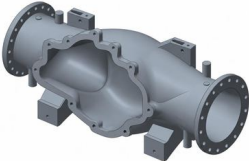
特征识别任务的测试样本采用具有明确B-rep几何、且可人工确认加工特征的B-rep模型。样本覆盖不同类型、不同数量的加工特征，并涵盖特征相交、嵌套等特殊情况。

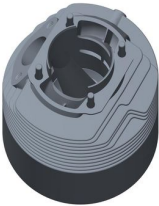
样本复杂度主要依据特征数量、特征边界完整程度、特征子类型判别难度，以及不同特征之间的相交、嵌套等相互关系。复杂度不单纯由零件面数或特征数量决定，具体如表3-2所示。基于上述难度分级，总共选取简单、中等和复杂三个级别的样本各2个。具体信息见表3-3。

表3-2 特征识别任务样本复杂度分级表

样本复杂度	选取标准	典型表现
简单	包含少量基础加工特征，特征边界基本完整、彼此独立，高层级类别和具体子类型能够由局部直接判断。	单个或少量独立特征；无明显特征间相互关系；特征边界完整。
中等	包含多个加工特征，需要结合多个几何区域进行区分；存在一定特征间相互关系，或具体子类型需要结合周围结构才能确定。	多种特征并存；轻度相交、局部截断或嵌套；部分边界受到其他特征影响。
复杂	存在多层嵌套、严重相交或复合特征，部分原始特征边界被其他特征截断、覆盖或共享，使特征发现、分解和细分类均存在较大困难。	多个特征间两两相交，多重特征嵌套。

表3-3 特征识别任务样本清单

样本ID	CAD模型渲染图	复杂度	复杂度归类说明
FR-01		简单	特征数量较少，孔和凹槽等特征相对独立，边界清楚。
FR-02		简单	孔、槽等特征数量较多，但是轮廓清楚，基本相互独立，可以逐一识别。
FR-03		中等	侧向接口与主体型腔相接，交界处边界不完整，需结合不同方向判断。

样本ID	CAD模型渲染图	复杂度	复杂度归类说明
FR-04		中等	内腔与侧面开口连通，部分边界被打断，存在较多的特征相交情况。
FR-05		复杂	多层环形结构紧密叠置，内部特征被遮挡，难以分清各层边界。
FR-06		复杂	多部件相互遮挡，孔、槽等结构密集，存在多个特征相交和多重特征嵌套，难以确定特征边界及归属。

3.2.2 测试条件设计

(1) 主实验设计

主实验对全部测试样本使用统一的Prompt。该Prompt不提供预设的特征类型列表，要求GPT-6自主识别B-rep模型中的加工特征，逐特征给出高层级类别与具体子类型，并说明主要判断依据及特征间关系。主实验仅改变测试样本，以考察不同复杂度下的特征识别表现。具体Prompt内容为：“请识别所提供B-rep模型中加工特征。对于每一个独立的特征实例：1.先给出其所属的较高层级特征类别（比如孔，仅作为举例），再尽可能进一步识别其具体子类型(比如通孔，仅作为举例)；2.指出该特征在模型中的位置，并说明支持这一判断的主要几何或拓扑依据；3.如果不同特征之间存在相交、嵌套、包含、截断、重复或其他明显的相互关系，请一并指出。请系统检查整个零件，不要只识别最明显的特征。对于无法可靠细分的特征，请保留在能够确定的层级，并说明存在的不确定性”。

(2) 扩展实验设计

在主实验基础上，选取代表性复杂样本开展扩展实验，考察网络搜索权限的影响，关注GPT-6在缺少外部网络信息时的特征发现、层级判定与特征关系识别表现。

说明：主实验和扩展实验均在独立会话中测试，且在特征识别完成后统一追加可视化输出Prompt，用于生成逐特征高亮的HTML核查材料。该步骤并不修改或替换原始识别结果。

3.3 测试结果展示

本节分别展示主实验和扩展实验的结果。主实验关注GPT-6对加工特征的自主发现、层级分类、几何定位以及特征间相互关系的识别能力；扩展实验用于观察关闭网络搜索后，工作流程和识别结果是否发生明显变化。两类实验结果分别展示与评价。

3.3.1 主实验结果

本节首先汇总各测试样本的特征识别结果，对比不同复杂度下的表现；随后逐样本展示CAD模型、GPT-6识别出的代表性加工特征及其B-rep几何对应，并从特征发现与实例划分、类别与子类型识别及特征关系识别三个方面进行分析，重点关注相交、嵌套和边界截断情况下的表现。

(1) 总体概览

主实验的特征识别结果按特征发现与实例划分、类别与子类型识别、特征间关系识别三个维度评估，如表3-4所示。各测试样本的用时见表3-5。

表3-4 特征识别任务主实验结果汇总表

样本ID	复杂度	特征发现与实例划分	类别与子类型识别	特征间关系识别
FR-01	简单	正确	正确	正确
FR-02	简单	正确	正确	正确
FR-03	中等	基本正确， 部分非特征面组被误识别为特征	正确	正确
FR-04	中等	正确	正确	正确
FR-05	复杂	基本正确， 部分非特征面组被误识别为特征	正确	正确
FR-06	复杂	基本正确，存在少量识别错误	基本正确，存在少量特征类别识别错误	正确

表3-5 特征识别主实验各样本测试用时

样本ID	FR-01	FR-02	FR-03	FR-04	FR-05	FR-06
用时	24分钟	37分钟	38分钟	43分钟	49分钟	2时1分钟

(2) 用例分析

FR-01（简单）

FR-01的代表性特征识别结果如图3-1所示。图3-1(b)中，深盲腔、腔口闭合环槽和外扩口沿被划分为不同实例；环槽与深腔之间由实体环带分隔，二者为相邻而非直接相交关系。图3-1(c)中，角部凸台、台阶通孔和锥底盲孔得到区分，台阶通孔的大、小径段归为一个复合孔实例；受截断的局部圆角仍归入同一过渡特征。图3-1(d)中，上、中、下三处侧壁通孔分别成组，并与内部主腔连通。

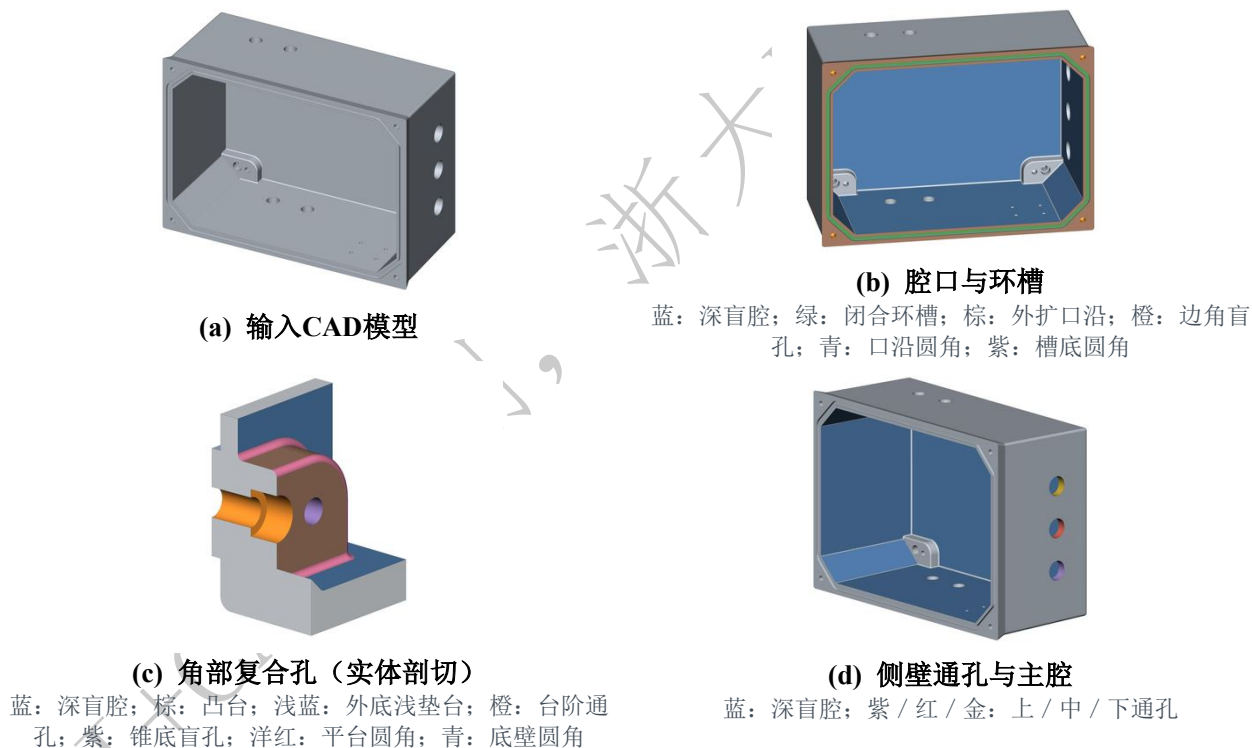


图3-1 FR-01代表性特征识别结果图

FR-02（简单）

FR-02的代表性特征识别结果如图3-2所示。图3-2(b)中，蓝色六处异形平底型腔环绕棕色中央凸台，虽共用凸台侧壁，却各有独立底面，因而被划分为六个实例；绿色窄长圆端槽、橙色凸台内通孔与紫色角部开口沉台也得到区分。图3-2(c)中，蓝、绿两处锥底盲孔沿同一轴线相向延伸，中间的灰色实体将孔尖隔开，表明它们是两个独立盲孔；橙色小盲孔另成实例。图3-2(d)中，蓝色中央盲腔底部嵌有绿色环状加深凹区；橙色四处两级型

腔分别将浅层台肩和深层凹底归为同一复合实例，紫色浅型腔和棕色角部开口沉台则独立成组。

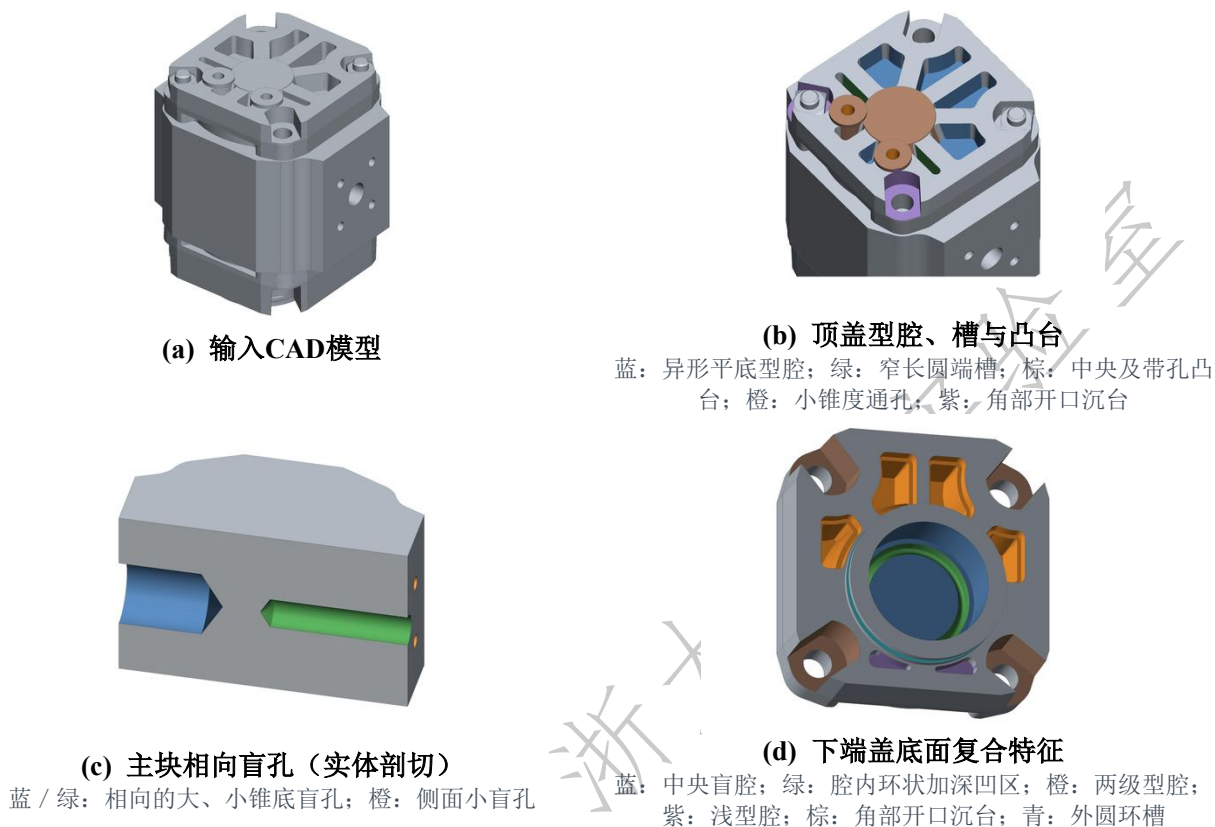


图3-2 FR-02代表性特征识别结果图

FR-03（中等）

FR-03的代表性特征识别结果如图3-3所示。图3-3(b)中，洋红色区域被输出为“整体空心壳体”主体特征；它描述的是零件基体，而非可独立计数的加工特征，体现表3-4所列的非加工特征误计。棕色区域为中分结合凸缘。图3-3(c)中，蓝色外包曲面壳腔与绿色中央弯曲流道受到棕色曲面隔壁的局部分隔，并在半环开口内侧相通。图3-3(d)中，蓝色半圆柱轴向座内嵌橙色双斜肩扩径槽；绿色半环台阶座内嵌紫色矩形截面半环槽，两组结构均在中分面形成开边。

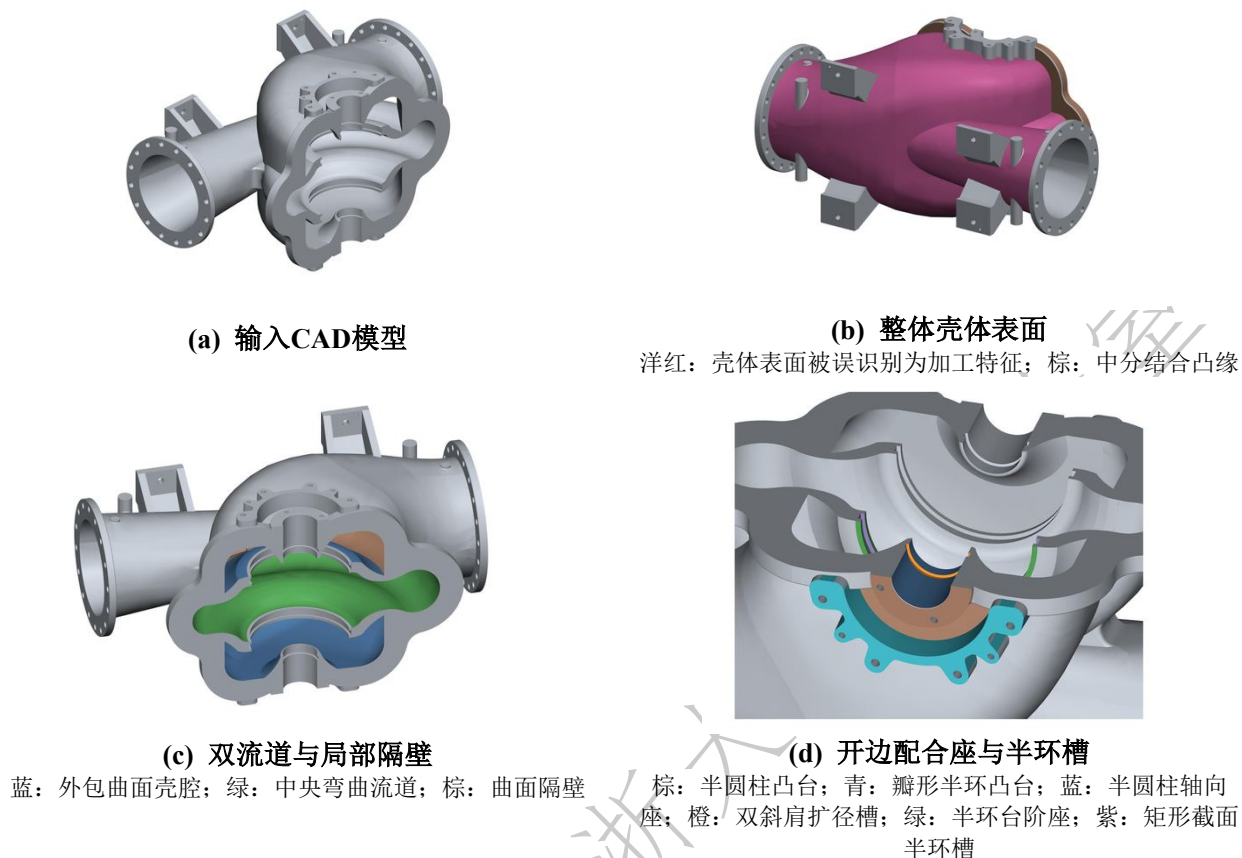
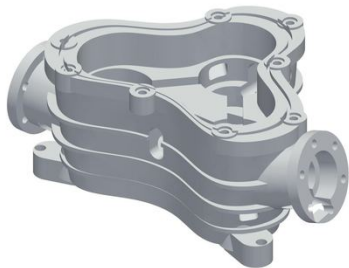


图3-3 FR-03代表性特征识别结果图

FR-04（中等）

FR-04的代表性特征识别结果如图3-4所示。图3-4(b)展示主腔内部三组重复的嵌套腔槽结构：浅蓝色主腔之下，蓝、绿、橙、紫四层特征沿深度方向逐级向内收缩，能够直观看到由浅到深的层次关系，以及局部被周围通道裁剪后的实际边界。图3-4(c)突出顶部一层的圆孔、对应凸台和异形沟槽：蓝色小孔分布在顶缘一周，绿色凸台与之对应，棕色沟槽沿顶缘连续绕行，并在局部与孔口和凸台相交，体现出这一层特征之间较强的组合关系。图3-4(d)中，蓝色阶梯接口孔与青色横向通道在侧接口处相交。法兰上四个紫色安装通孔保持完整圆周，另外两个橙色安装孔被横向通道截断，体现了复杂相交条件下的孔特征识别结果。

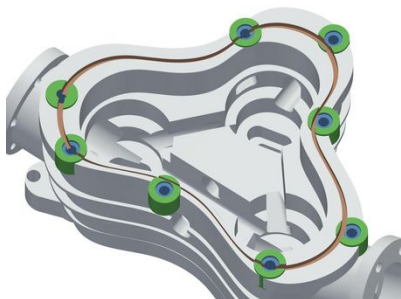


(a) 输入CAD模型



(b) 四级嵌套腔槽

浅蓝：主腔；蓝：一级圆形沉腔；绿：二级椭圆形平底沉腔；橙：三级长圆槽；紫：四级平底盲长圆槽



(c) 顶缘圆孔与异形沟槽

蓝：顶缘小孔；绿：对应凸台；棕：顶部异形沟槽



(d) 侧接口与受截断的安装孔

棕：侧接口法兰；蓝：阶梯接口孔；青：横向通道；紫：完整安装通孔；橙：两个被横向通道截断的安装孔

图3-4 FR-04代表性特征识别结果图

FR-05（复杂）

FR-05的代表性特征识别结果如图3-5所示。图3-5(b)中，洋红色区域对应上盖球面拱形基体与圆盘基部等整体外形，它们属于上盖基体，而非可独立计数的加工特征，反映了FR-05存在非加工特征实例误计的问题。图3-5(c)中，蓝色主球面凹腔内嵌绿色次级球面凹区，紫色过渡圆角连接两者，橙色斜通孔从次级凹区连通外部；青色环形定位止口外侧，玫红色下端波状轮廓安装法兰围绕四个棕色上盖安装通孔，展示了内腔嵌套及周边安装结构。图3-5(d)中，蓝色周向散热片分层排列，棕色径向孔座包围橙色平底盲孔，绿色上端波状接合法兰位于上方；孔座与相邻片层相交并截断局部片缘，体现了外部复杂特征与局部孔特征的组合关系。

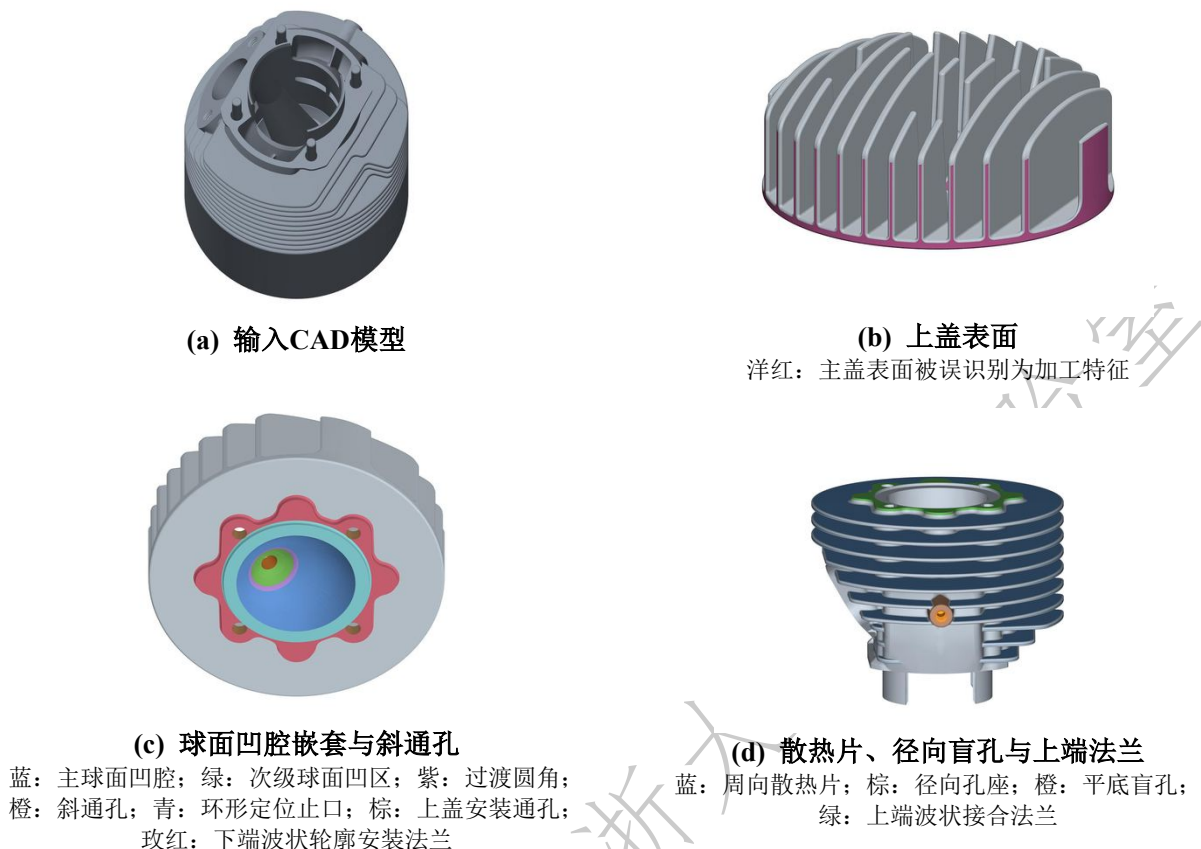


图3-5 FR-05代表性特征识别结果图

FR-06（复杂）

FR-06的代表性特征识别结果如图3-6所示。图3-6(b) 采用与HTML视图一致的局部剖切方式展示实例划分错误，可以看到两个圆孔在空间上彼此隔离，但GPT-6仍将其识别为同一个孔特征；图3-6(c) 同样采用局部剖切视图展示类别错误，被标出的圆柱面实际位于实体外侧，却被误识别为孔类特征。图3-6(d) 展示散热片、端面散热肋、端部凹区 / 筋间开口、截断横向孔壁共存的复杂区域。

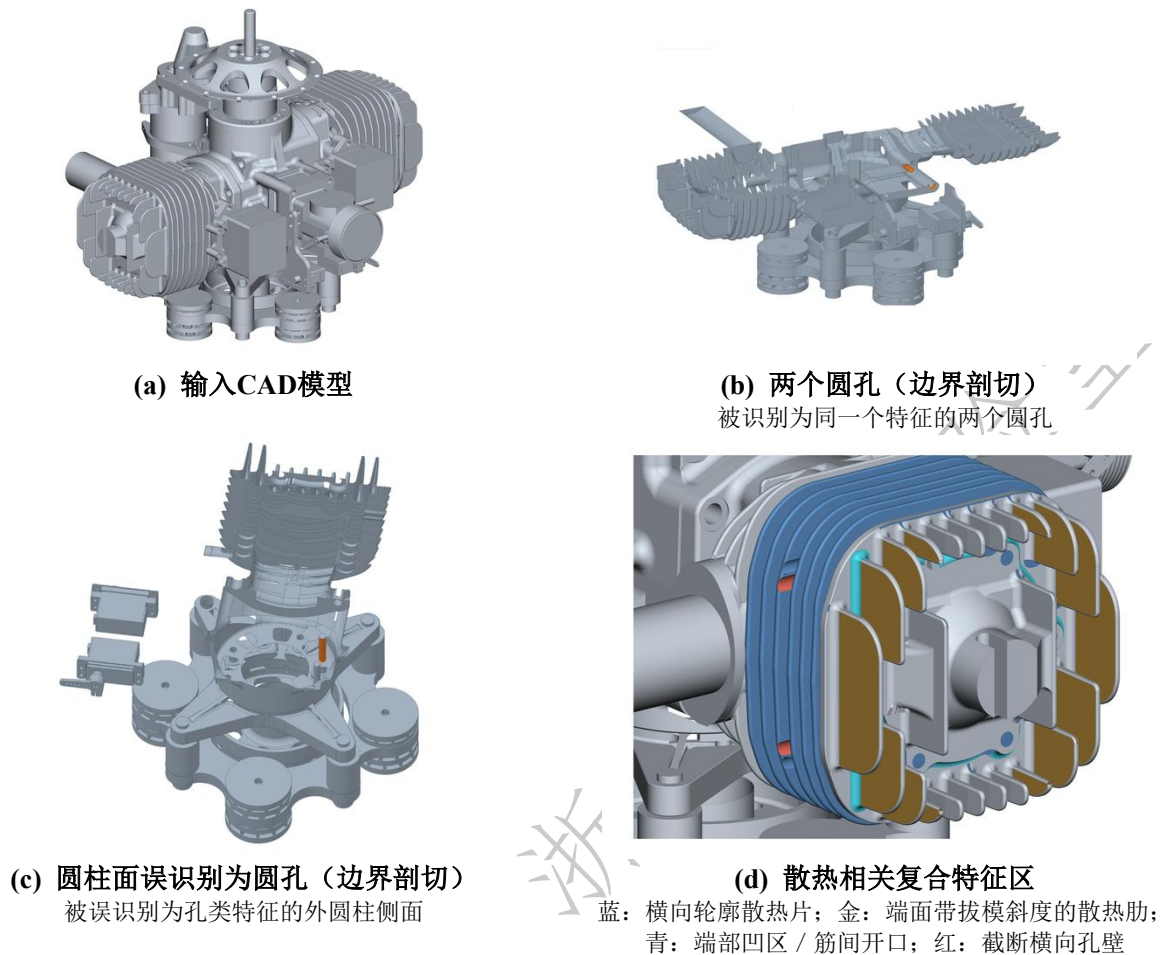


图3-6 FR-06代表性特征识别结果图

3.3.2 扩展实验结果

扩展实验仅选取复杂样本FR-06，比较开启与关闭网络搜索时的特征识别结果。以下对同一几何区域采用相同视角并列展示：左侧为有网络，右侧为无网络；彩色区域表示相应结果识别到的特征，灰色仅为对照用的模型表面，不直接代表无特征。重点比较复杂区域的特征类别、实例划分与局部几何定位。无网络搜索实验的测试用时见表3-6。

表3-6 特征识别无网络搜索实验测试用时

样本ID	无网络_FR-06
用时	47分钟

(1) 上部轮盘：窗口及不同孔口的识别基本一致

FR-06上部轮盘的有、无网络识别结果如图3-7所示。轮盘上的扇形窗口、内外两圈锥形孔口与中央台阶轮毂相互邻近，需区分各自的几何边界。

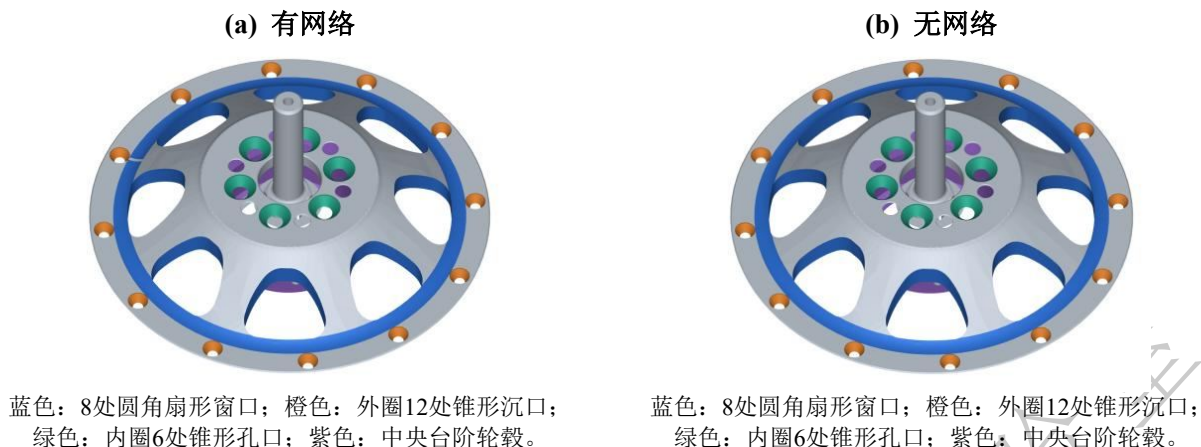


图3-7 FR-06上部轮盘特征有网络/无网络同区域特征识别对照

对比分析：图3-7(a)、(b)中，蓝色扇形窗口均沿轮盘周向逐一分组，橙色外圈沉口与绿色内圈孔口分别着色，紫色轴毂也被单独识别。两种条件下这些特征的类型、数量及对应区域基本一致，未见明显的窗口合并、孔口混淆或位置偏移。

(2) 夹紧连杆：有网络时进一步识别出通孔子类型

夹紧连杆上的4处小孔沿杆身一侧均匀分布，局部孔壁与两端孔口清晰可见，适合比较两种条件在孔类子类型判断上的差异。图3-8给出了同一局部区域的有、无网络识别结果。

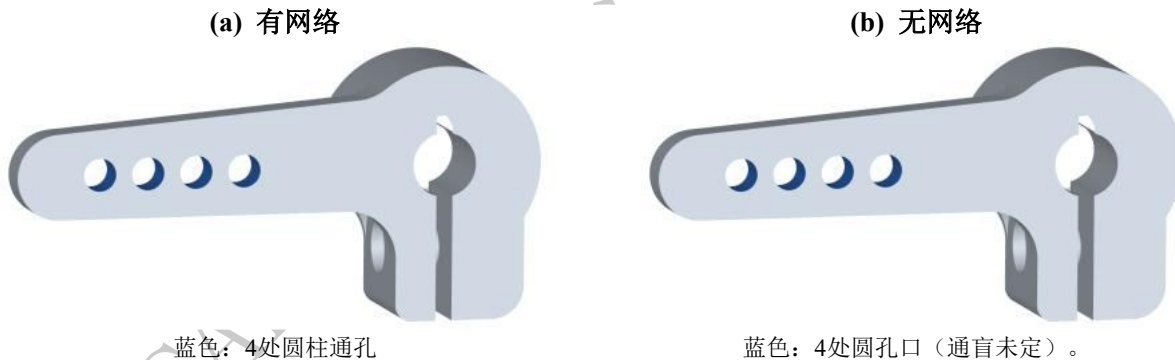


图3-8 FR-06夹紧连杆四处小孔有网络/无网络同区域特征识别对照

对比分析：图3-8(a)、(b)中，两种条件均定位到相同的4处小孔孔壁，局部几何位置一致。差异在于，有网络结果进一步将其识别为圆柱通孔；无网络结果虽然发现了相同孔位，但仅停留在圆孔层级，未继续完成通盲子类型判断。在该局部区域，有网络结果在孔类细分上更完整。

(3) 台阶轴毂：无网络时正确区分平底盲孔与未定孔段

台阶轴毂侧面具有两处邻近的径向孔，一处孔壁内端连接平面孔底，另一处仅保留圆柱孔段。图3-9放大同一侧面，比较两种条件是否利用孔底完成通盲细分。

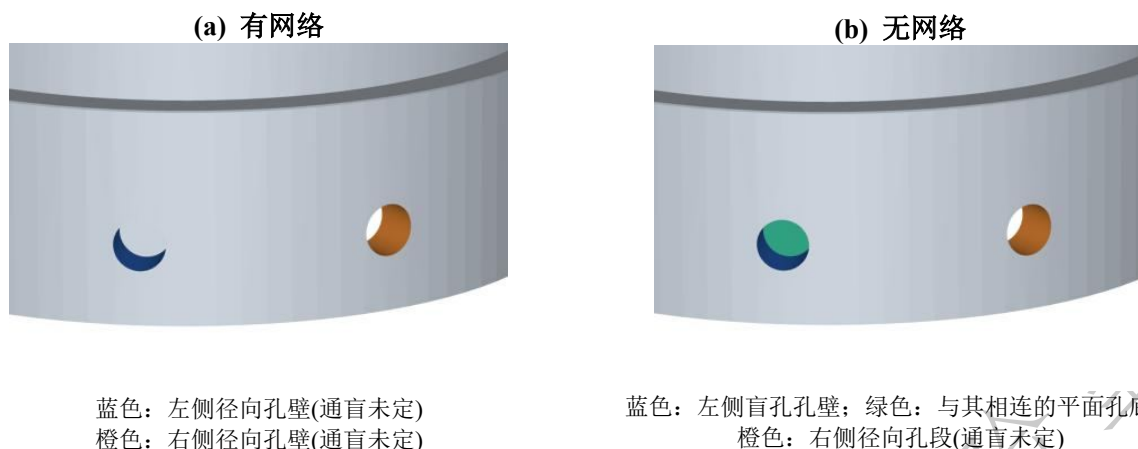


图3-9 FR-06轴毂侧面径向孔有网络/无网络同区域特征识别对照

对比分析：图3-9(a)中，有网络结果发现两处径向孔壁，但均停留在孔这一层级。图3-9(b)中，无网络结果将左侧蓝色孔壁和绿色平面孔底组合判断为平底盲孔，同时保留右侧橙色孔段的通盲不确定性。由图可见，无网络结果在该处利用孔底信息完成了进一步细分；有网络结果虽识别到孔的大类，但未继续识别为盲孔。

(4) 扩展结果分析

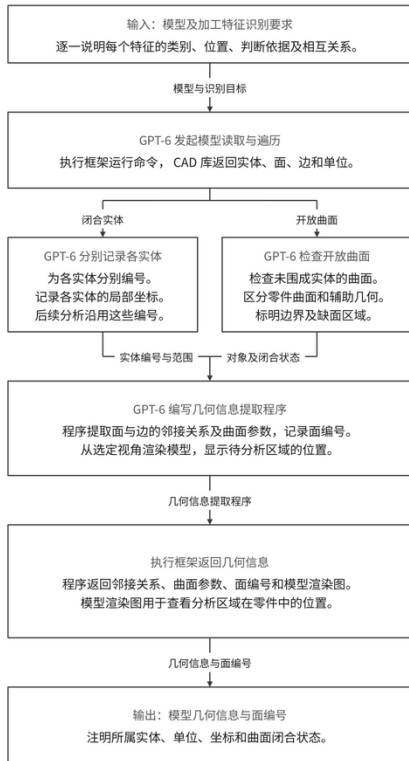
三组对照测试呈现出不同情况：轮盘窗口和孔口识别基本一致；夹紧连杆区域中，有网络结果进一步完成了4处小孔的通孔子类型判断；台阶轴毂则在无网络条件下进一步识别出平底盲孔。说明GPT-6在有/无网络条件下的特征识别能力差距并不大，其性能表现更多来自于对CAD内核工具/代码工具的调用和对图像的理解（详见3.4节的流程分析）。此外，通过本组扩展实验还能发现GPT-6在复杂实体模型上的特征识别并不充分，存在部分特征实例识别结果比较粗糙的情况（如图3-8和图3-9）。

3.4 GPT-6推理过程分析

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对特征识别任务可能采用如下推理流程：

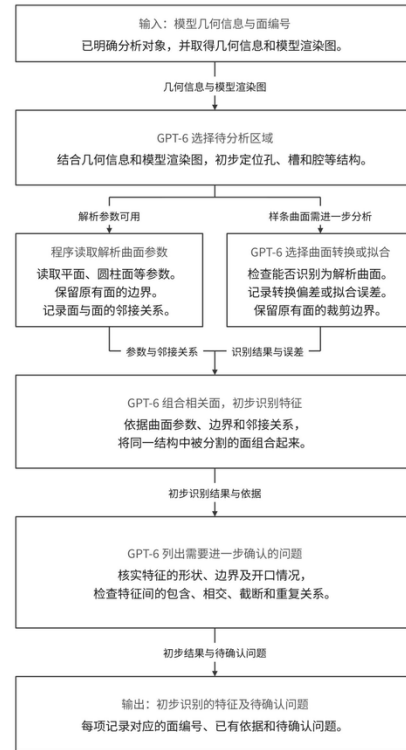
① 模型读取与信息提取

区分实体和开放曲面，整理模型的几何信息



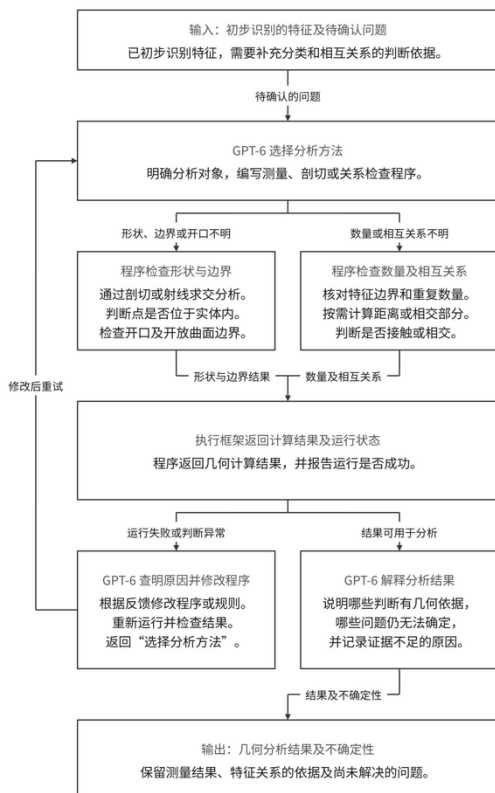
② 初步识别与问题梳理

根据几何信息初步识别特征，列出需要确认的问题



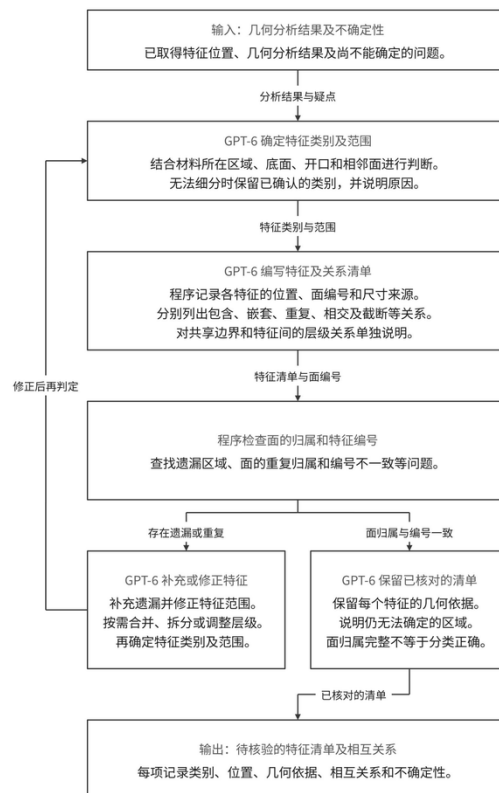
③ 几何分析与补充验证

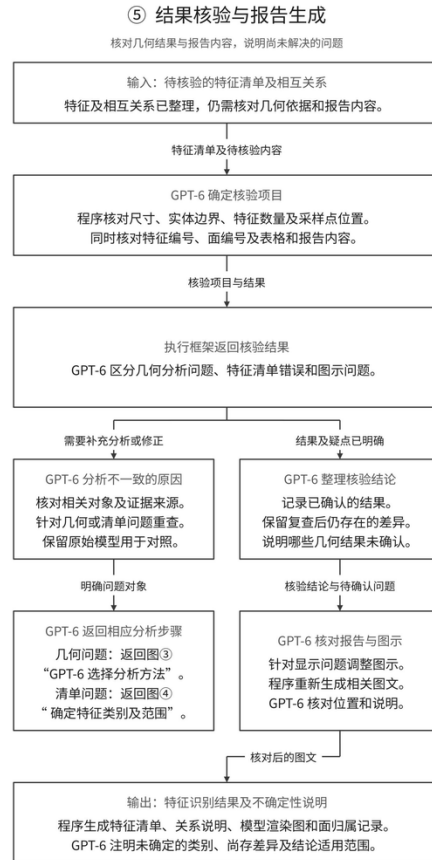
针对待确认问题开展分析，根据结果调整方法



④ 特征分类与结果整理

确定各个特征的类型、范围及相互关系





3.5 能力边界

在特征识别任务上，GPT-6已能够结合曲面参数、面邻接关系、必要的剖切分析及实体内外检查，对孔、槽、型腔、凸台和过渡结构进行识别，并将识别结果对应到具体B-rep面。对于边界清晰、彼此分离的加工特征，GPT-6能够正确识别其类型，并进一步辨认其局部结构和层次关系。具体来说，对于孔类特征，能够区分通孔、锥底盲孔、台阶孔等不同类型；对于型腔类特征，则能够识别不同层级的型腔，并结合底面位置和连通关系区分不同的特征。

当前的主要能力边界集中在存在多层嵌套、相交或边界截断的复杂区域的特征识别上。例如，GPT-6有时会将空间上分离的两个孔错误识别为同一个特征，或将圆柱面误判为圆孔；对于部分已经定位到的孔特征，也未能充分结合孔底和开口信息进一步判断其具体子类型（通孔/盲孔）。结合3.4节的任务流程分析可以看到，GPT-6在部分判断过程中会使用模型渲染图辅助理解几何结构。对于视角受限、遮挡明显或局部结构在渲染图中不易辨认的区域，图像信息可能会对特征面组选取和特征类型判断产生干扰，这可能是上述细节误判的原因之一。

总体来看，GPT-6已经能够较可靠地识别基础特征，并处理部分复杂特征组合，但在复杂模型中，对全部加工特征进行识别的结果仍不够稳定。特别是在特征边界不清晰、多个特征相互影响，或局部结构难以从模型视图中直接辨认时，更容易出现多个特征被误识别成一个特征或特征细分类型不清等问题。

AI+CAD评测，浙大DMA实验室

4 工程图到CAD模型重建/生成

4.1 任务目标与评价标准

工程图到CAD模型重建/生成任务主要测试GPT-6对机械工程图的理解与三维CAD模型重建能力。测试重点关注GPT-6能否正确理解工程图中的多视图投影、尺寸标注及局部几何特征，并据此生成几何拓扑准确、关键尺寸正确的三维CAD模型。基础测试覆盖单个零件以及装配体中的多零件建模场景。在此基础上，通过工程图尺寸或联网条件变化等补充测试，进一步考察模型在不同输入条件下的表现。

表4-1 工程图到CAD任务说明

测试要素	说明
测试问题	GPT-6能否根据工程图正确理解零件的三维结构，并生成相应的CAD模型？
核心输入	包含多视图、尺寸标注及必要标注的工程图
核心输出	三维CAD模型
评价标准	生成结果是否为几何、拓扑合法且一致的CAD模型（实体有效性），并检查零件结构是否与工程图一致、图中标注尺寸是否得到准确还原；对于多零件用例，进一步验证各零件能否按预期完成装配。

4.2 测试用例设计

4.2.1 测试样本选取

工程图到CAD模型重建/生成测试样本主要选取公开机械制图及CAD教学资料中的工程图纸。样本选取优先考虑图纸信息完整、形状具有代表性，且能够获得对应三维模型或其他可靠参考结果的案例。同时，为覆盖不同难度的建模任务，样本在结构复杂程度、工程图信息量以及零件数量等方面保持一定差异，并同时包含单零件和多零件建模场景。

具体而言，样本复杂度综合考虑工程图信息量、零件结构复杂程度以及建模关系复杂程度进行划分。其中，多零件样本还进一步考虑零件数量及装配关系。具体划分标准如表4-2所示。

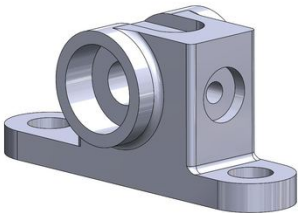
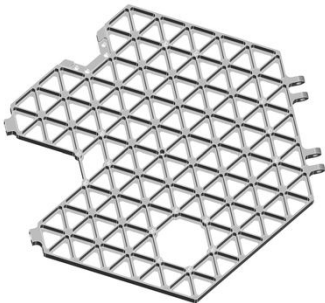
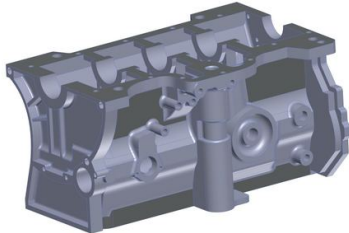
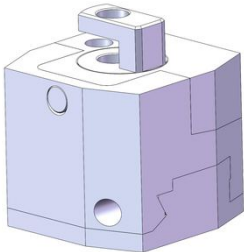
表4-2 工程图到CAD模型重建/生成任务样本复杂度分级表

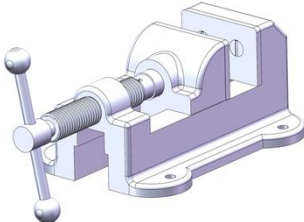
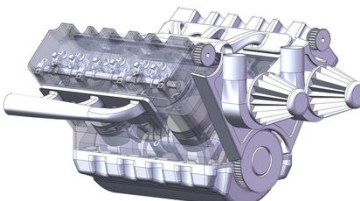
样本复杂度	选取标准	典型表现
简单	工程图中几何数量和视图数量较少，零件主要结构可由基本投影视图和少量剖视直接确定，跨视图关联和内部结构推断需求较低。	主要根据少量视图即可恢复零件整体结构，局部特征数量较少、关系明确。
中等	视图信息量明显增加，需要进行一定的跨视图关联；或零件中包含较多局部细节、曲线关系及重复特征，但主体仍具有较强规律性；多零件任务中，零件数量较少且装配关系较明确，但需要进行跨零件信息关联。	需要综合多个视图或多个零件的信息完成建模，但仍可利用明显的形状规律、重复关系或较明确的装配关系降低推理难度。
复杂	视图信息密集，包含较多投影视图、剖视或局部视图，结构细节类型丰富且缺乏单一重复规律，需要综合大量分散信息确定完整几何；多零件样本还可能包含较多零件及需要结合其他零件信息判断的局部结构。	需要综合大量分散的几何与尺寸信息，并处理复杂的跨视图关系；在多零件场景下，部分结构还可能需结合零件间的对应关系进行理解。

依据上述标准，主测试共选取8个样本，其中包括5个单零件样本和3个多零件样本。单零件样本包含2个简单样本、2个中等样本和1个复杂样本；多零件样本包含1个中等样本和2个复杂样本。具体样本清单如表4-3所示。

表4-3 工程图到CAD模型重建/生成任务主测试样本清单

样本ID	CAD模型渲染图	样本类型	复杂度	复杂度归类说明
D2C-P01		单零件	简单	单张工程图包含三视图及1个剖视图，视图数量较少，主要几何结构及内部关系可由基本投影视图直接确定，跨视图关联需求较低。

样本ID	CAD模型渲染图	样本类型	复杂度	复杂度归类说明
D2C-P02		单零件	简单	单张工程图包含三视图及2个剖视图，整体结构由常规孔、台阶及凸台等基本特征构成，主要几何关系可由少量视图直接恢复。
D2C-P03		单零件	中等	共由3张工程图、12个视图描述，输入信息量较大，并包含局部开孔等差异化细节；但主体结构以规则阵列为主，大量几何可通过重复规律确定，因此归为中等复杂度。
D2C-P04		单零件	中等	共由4张工程图、8个视图描述，包含较多曲线轮廓及相互交汇关系，需要结合不同视图确定局部边界和几何对应关系；但整体信息规模及结构层次仍明显低于复杂样本。
D2C-P05		单零件	复杂	由11个视图描述，包含大量内外部结构、孔系、凸台、腔体及局部细节，不同结构分散在多个投影视图和剖视信息中，且缺乏单一重复规律，需要综合大量跨视图信息完成完整建模。
D2C-A01		多零件	中等	由4个独立零件组成，各零件自身结构及装配关系相对简单，但需要分别完成多个零件建模，并结合跨零件尺寸及配合关系完成装配，因此归为中等复杂度。

样本ID	CAD模型渲染图	样本类型	复杂度	复杂度归类说明
D2C-A02		多零件	复杂	由11个零件组成，除零件数量明显增加外，还包含销钉配合、螺纹配合等不同装配关系，需要同时处理多零件几何建模、零件对应及多种配合关系。
D2C-A03		多零件	复杂	包含28项零件，零件数量和类型较多，内部结构复杂，单件建模中包含抽壳、放样等多种实际工程建模操作，同时存在大量内部装配及零件间关系，整体信息规模和建模链条均明显高于其他样本。

4.2.2 测试条件设计

(1) 主实验设计

主实验使用统一处理后的工程图样本和统一Prompt对全部测试样本进行测试。测试前统一遮盖网站名称、标识、专有型号及产品编号等可能暴露来源的信息，并采用中性化文件名；对于多零件样本中用于关联零件图与装配图的编号信息，则保留其对应关系。

主实验中，同类建模任务采用统一Prompt，不针对样本复杂度及具体结构提供额外提示，要求GPT-6根据输入工程图完成三维CAD模型（具体为B-rep模型）构建并输出相应模型文件及预览结果。对于多零件样本，在各零件建模完成后，进一步按照统一要求进行装配验证。主实验保持网络搜索权限开启，以考察GPT-6在正常使用条件下的工程图到CAD模型重建/生成能力。具体Prompt内容为：“根据提供的工程图，生成图中零件的三维CAD模型，并导出STEP文件。”（单零件），“根据提供的装配体各零件工程图，分别生成对应零件的三维CAD模型，并导出STEP文件。”（多零件）。

(2) 扩展实验设计

在主实验基础上，选取具有代表性的样本开展扩展实验，进一步考察工程图输入信息和网络搜索条件变化对重建/生成结果的影响。扩展实验主要设置两种条件变化：删除标注只给出零件包围盒信息、关闭网络搜索权限。各扩展实验均沿用主实验中的对应Prompt，

不针对具体样本增加额外提示；除所考察条件外，其余测试环境与输入处理方式保持不变。扩展实验结果分别与对应样本的主实验结果进行对照。具体设置如表4-4所示。

表4-4 工程图到CAD模型重建/生成任务扩展实验设置

设置	样本数据处理	网络搜索	测试目的
S1尺寸信息缺失	仅保留总长、总宽、总高、最大外径等整体外形尺寸，删除其余局部尺寸	开启	考察局部尺寸信息缺失时模型对零件结构、局部比例及多零件配合关系的推断能力
S2无网络条件	样本内容与主实验保持一致，不进行额外处理	关闭	考察无法获取外部网络信息时模型的建模表现

4.3 测试结果展示

本节分别展示主实验与扩展实验结果。主实验首先汇总全部测试样本在实体有效性、结构还原、尺寸准确及装配效果等方面的总体表现，随后按用例展示样例描述、生成结果及分析。扩展实验按照4.2.2节设置的测试条件逐项展示，并与相同样本在主实验中的结果进行对比。

4.3.1 主实验结果

(1) 总体概览

主实验共包含5个单零件样本和3个多零件样本。各样本的总体结果汇总如表4-5所示。各测试样本的用时见表4-6。

表4-5 工程图到CAD模型重建/生成任务主实验结果概览

样本ID	难度	实体有效性	结构还原	尺寸准确	装配效果
D2C-P01	简单	通过	丢失两个小孔	其他地方通过	-
D2C-P02	简单	通过	通过	背部圆角有问题	-
D2C-P03	中等	0/1	通过	螺纹倒角有问题	-
D2C-P04	中等	通过	型腔侧壁有问题	因受结构影响不通过	-
D2C-P05	复杂	通过	结构大体正确	因受结构影响不通过	-
D2C-A01	中等	通过	结构几乎完全正确	存在不影响的小问题	通过

样本ID	难度	实体有效性	结构还原	尺寸准确	装配效果
D2C-A02	复杂	通过	结构几乎完全正确	部分尺寸不匹配	装配产生重叠
D2C-A03	复杂	25/28	出现一定程度缺失	部分尺寸不匹配	无法进行装配

表4-6 工程图到CAD主实验各样本测试用时

样本ID	D2C-P01	D2C-P02	D2C-P03	D2C-P04	D2C-P05	D2C-A01	D2C-A02	D2C-A03
用时	28分钟	25分钟	37分钟	1时8分钟	30分钟	26分钟	28分钟	36分钟

(2) 用例分析

以下按照样本编号逐一展示主实验结果。每个用例包括样例描述、结果展示和结果分析；多零件用例另展示由评测方完成的装配验证结果。

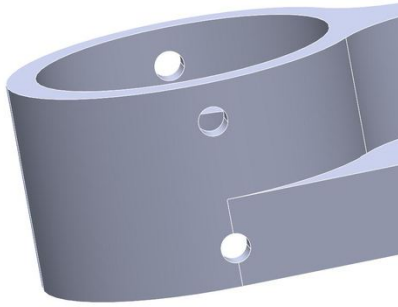
D2C-P01通用支架（简单）

结果展示

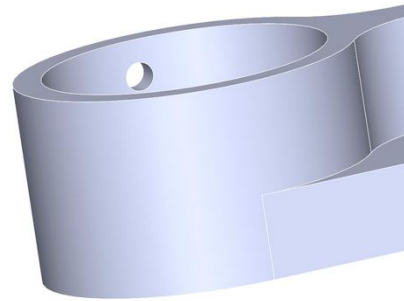


图4-1 D2C-P01工程图生成CAD结果

结果分析：生成的模型实体有效，结构尺寸与给出的工程图基本一致，右端圆柱凸台的阶梯孔尺寸基本正确，但其对称位置的径向孔缺失，如图4-2所示。



(a) 参考模型渲染图

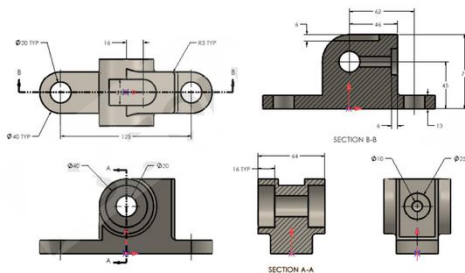


(b) 生成模型渲染图

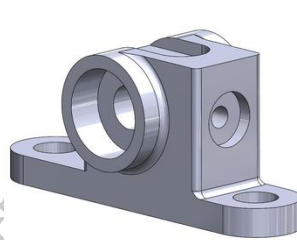
图4-2 D2C-P01工程图生成CAD细节

D2C-P02轴承支座（简单）

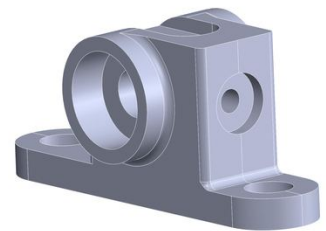
结果展示



(a) 工程图示例



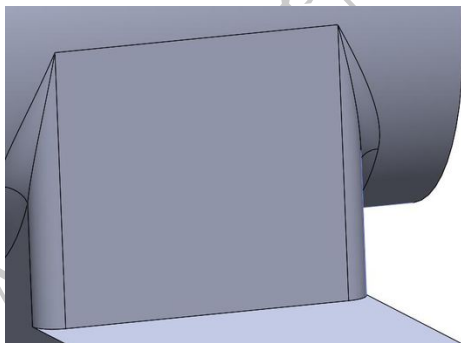
(b) 参考模型渲染图



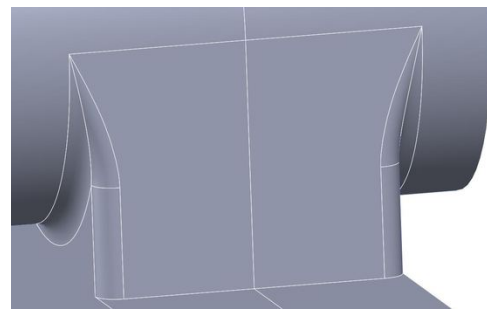
(c) 生成模型渲染图

图4-3 D2C-P02工程图生成CAD结果

结果分析：生成的模型实体有效，结构尺寸与给出的工程图基本一致，但背部竖直支承体两侧与圆柱结构之间的衔接与过渡轮廓不一致，如图4-4所示。



(a) 参考模型渲染图



(b) 生成模型渲染图

图4-4 D2C-P02工程图生成CAD细节

D2C-P03 TITAN-513M（中等）

结果展示

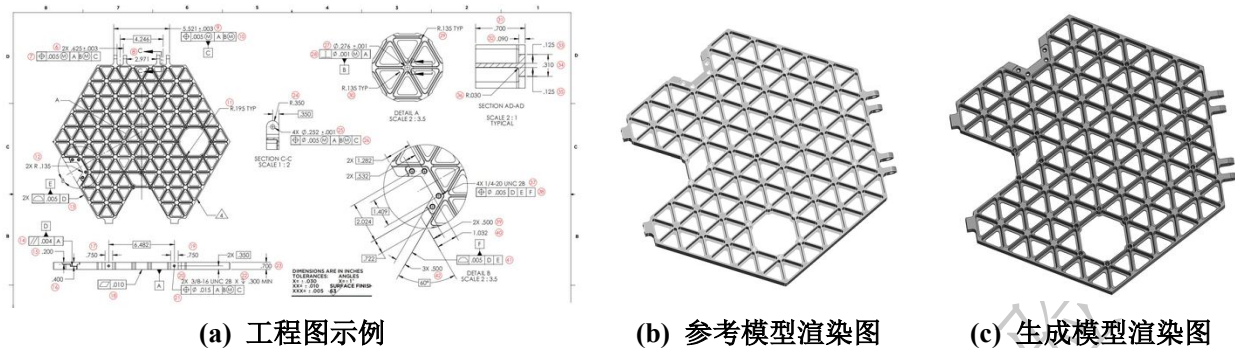


图4-5 D2C-P03工程图生成CAD结果

结果分析：生成模型实体出现不合法的面相交，整体结构与参考模型基本一致，规则阵列、局部开孔及主要几何关系均得到正确还原，结构还原检查通过。但存在如图4-6所示的螺纹孔入口处倒角处理与参考模型不一致，生成模型的倒角未正确过渡至螺纹小径，导致螺纹入口局部几何及相关尺寸参数存在偏差，进一步导致出现非法相交面。因此，该样本结构还原通过，但参数准确性检查未通过。

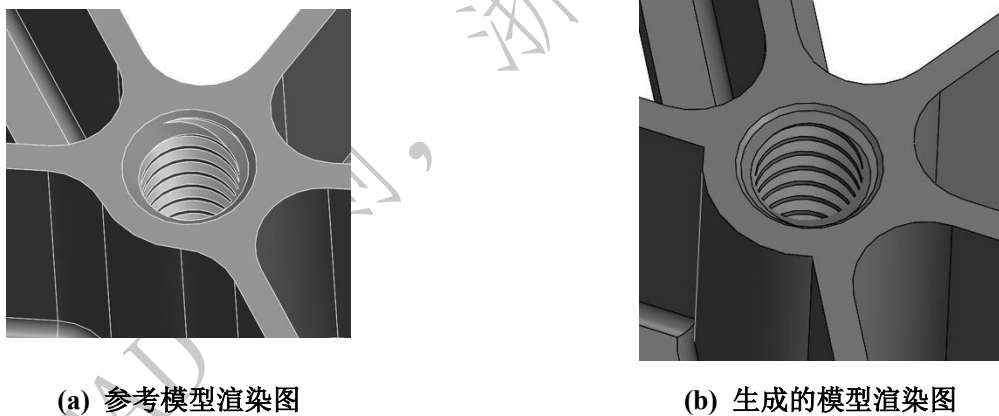


图4-6 D2C-P03工程图生成CAD细节

D2C-P04 TITAN-820（中等）

结果展示

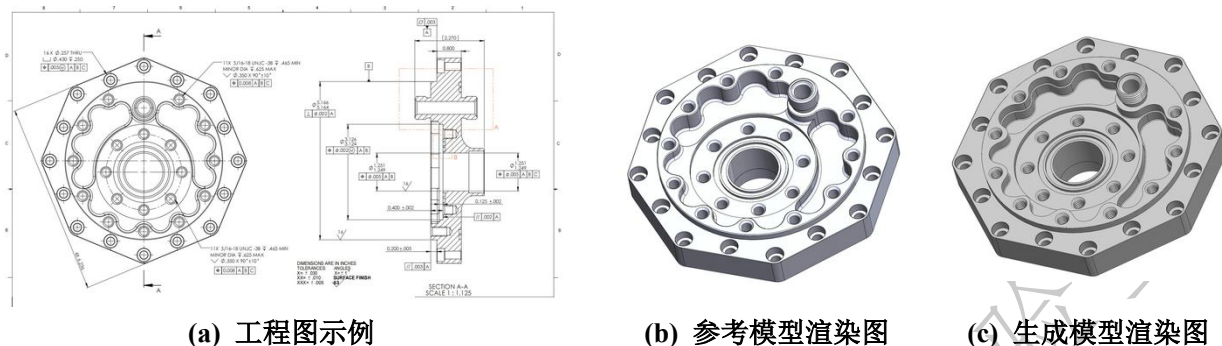


图4-7 D2C-P04工程图生成CAD结果

结果分析：生成模型实体有效，整体结构与参考模型基本一致，并能够根据工程图中的螺纹标注正确生成对应螺纹结构，说明其对局部加工特征的识别较为准确。但在零件内部侧壁结构上，未能正确理解剖视图所表达的层次关系。如图4-8所示，参考模型中该区域应依次形成“倒角—侧壁—圆角”的连续过渡，而生成模型将其错误简化为“两层侧壁—圆角”结构，导致内部轮廓及局部参数关系与参考模型不一致。因此，该样本主体结构还原较好，但内部复杂剖面结构的理解仍存在偏差。

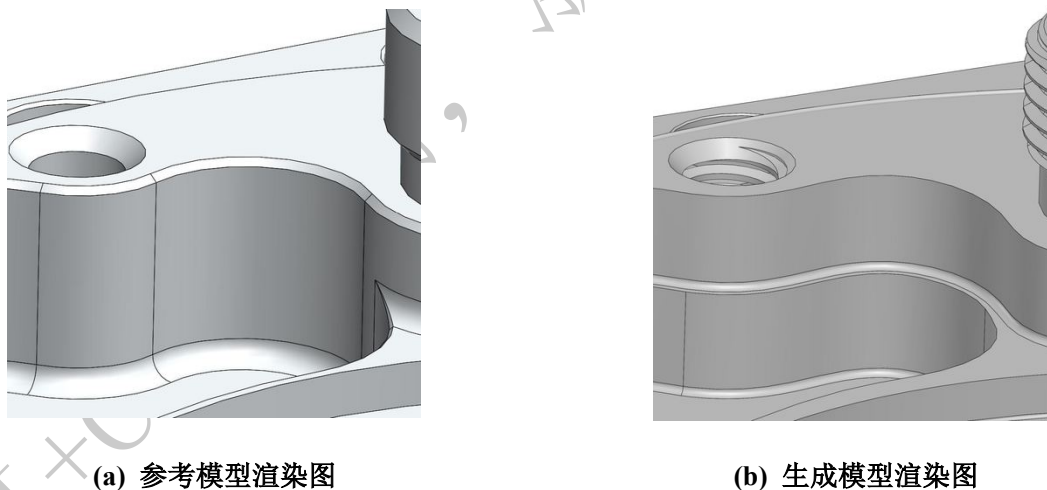


图4-8 D2C-P04工程图生成CAD细节

D2C-P05发动机缸体（复杂）

结果展示

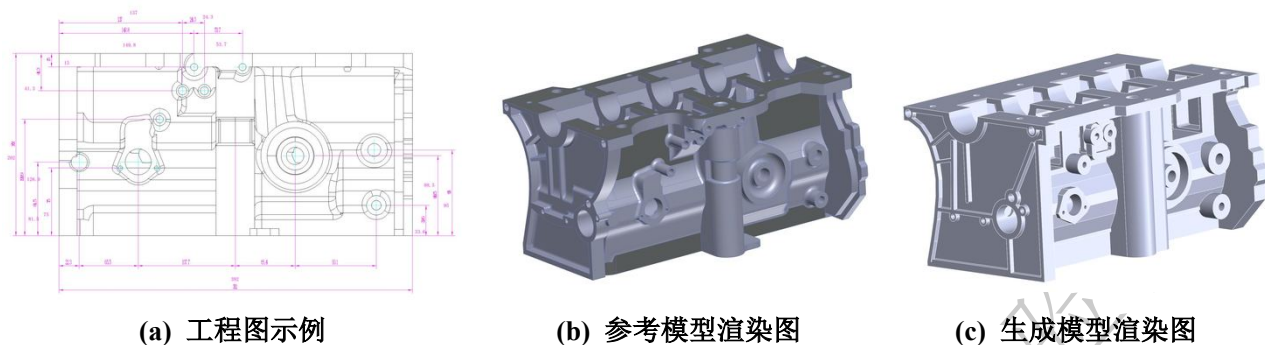


图4-9 D2C-P05工程图生成CAD结果

结果分析：生成模型实体有效，整体上较好还原了发动机缸体的主要外形、孔系、内部腔体及加强结构，主体结构与参考模型基本一致，但复杂局部结构及参数关系仍存在明显偏差。如图4-10所示，部分区域未正确形成“型腔内再开孔”的结构层次，中间腔体也出现了不合法的尖角轮廓；此外，部分原本应贯通的孔被错误建模为盲孔，导致局部拓扑关系发生变化。与此同时，多层台阶、腔体深度、孔深及局部壁厚等参数与参考模型仍存在一定差异。因此，该样本整体结构还原基本正确，但局部结构完整性及参数准确性检查未通过。

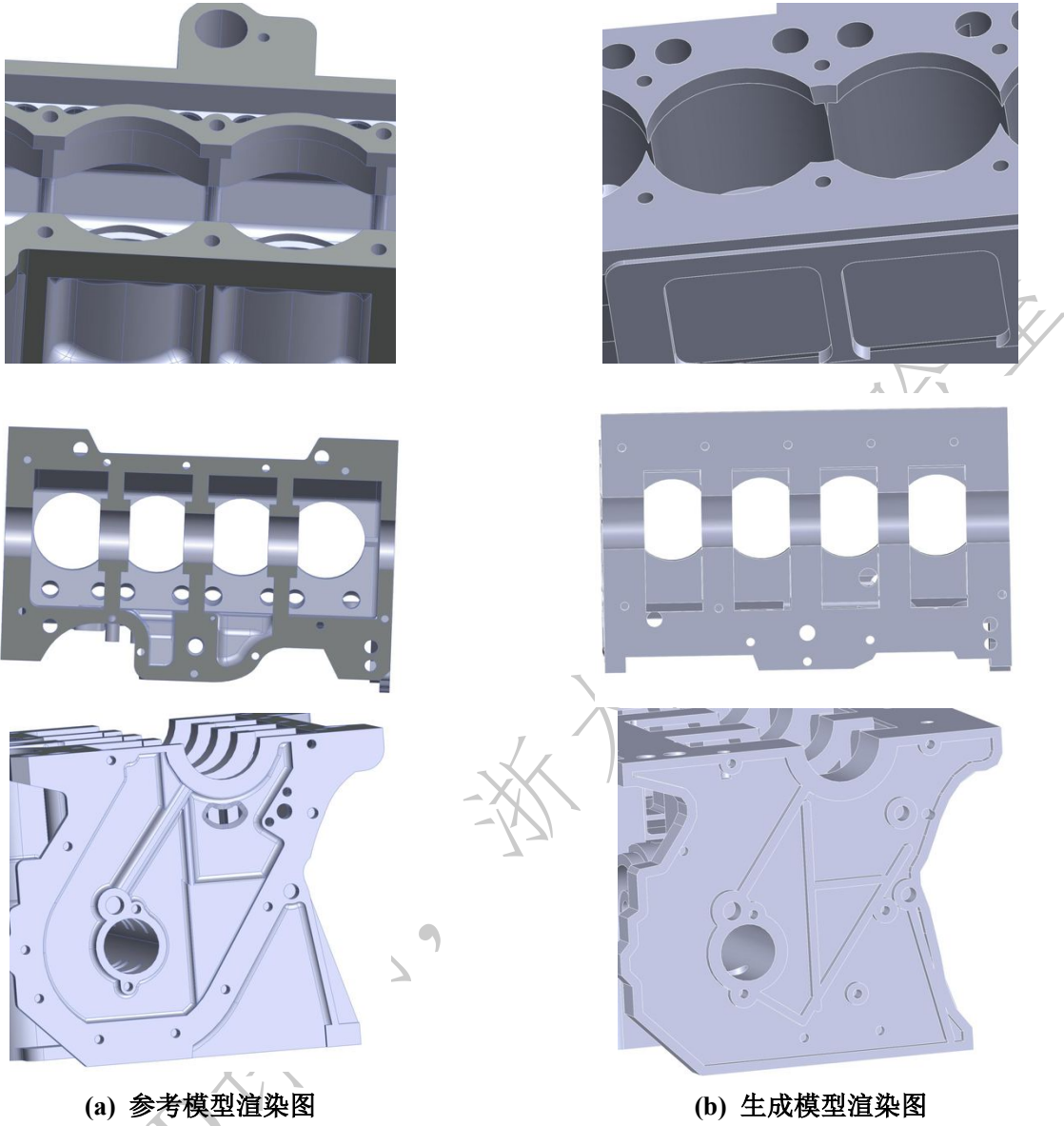
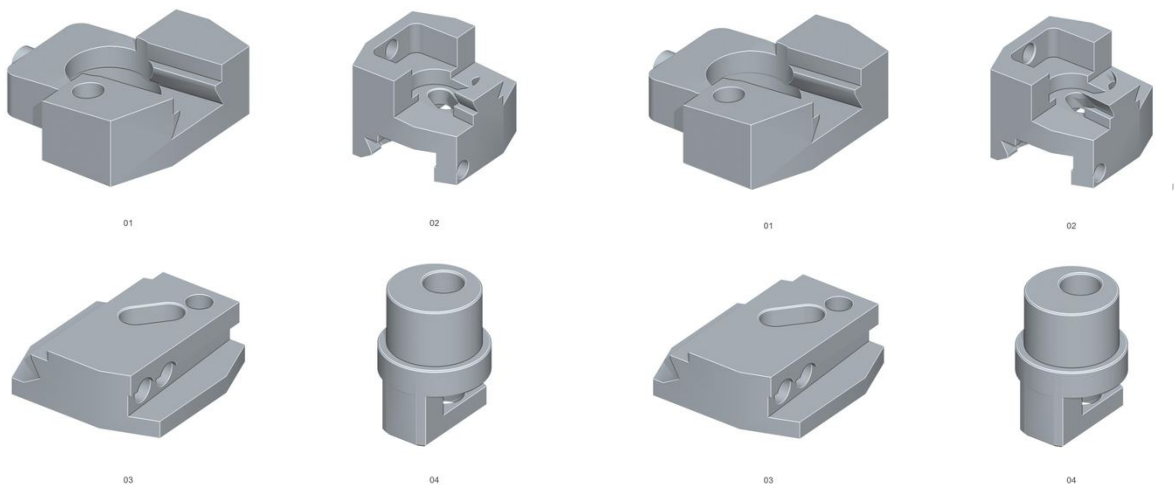


图4-10 D2C-P05工程图生成CAD细节

D2C-A01技能五轮铣削滑动装配件（中等）

结果展示

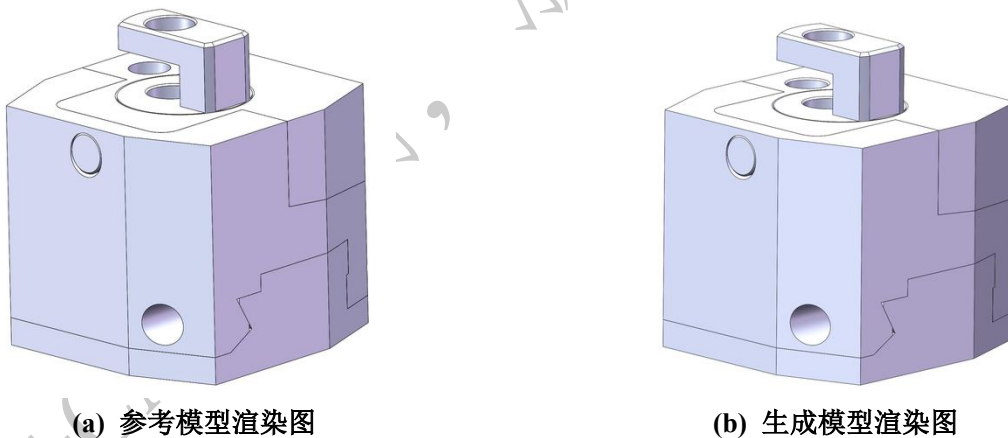


(a) 参考模型渲染图

(b) 生成模型渲染图

图4-11 D2C-A01工程图生成CAD结果

装配验证



(a) 参考模型渲染图

(b) 生成模型渲染图

图4-12 D2C-A01工程图生成CAD装配结果

结果分析：生成模型各零件实体有效，其主体结构、主要孔槽特征及相互配合关系被较完整地还原。装配后整体形态与参考模型基本一致，能够完成预期装配，结构还原整体正确。局部尺寸与几何参数仍存在少量偏差，如图4-13所示，参考模型中该处应通过斜面切削去除尖角，而生成模型未完全形成对应斜面，残留局部尖锐结构。该问题未对整体装配关系产生明显影响，但局部角度及尺寸参数的还原仍存在一定误差。

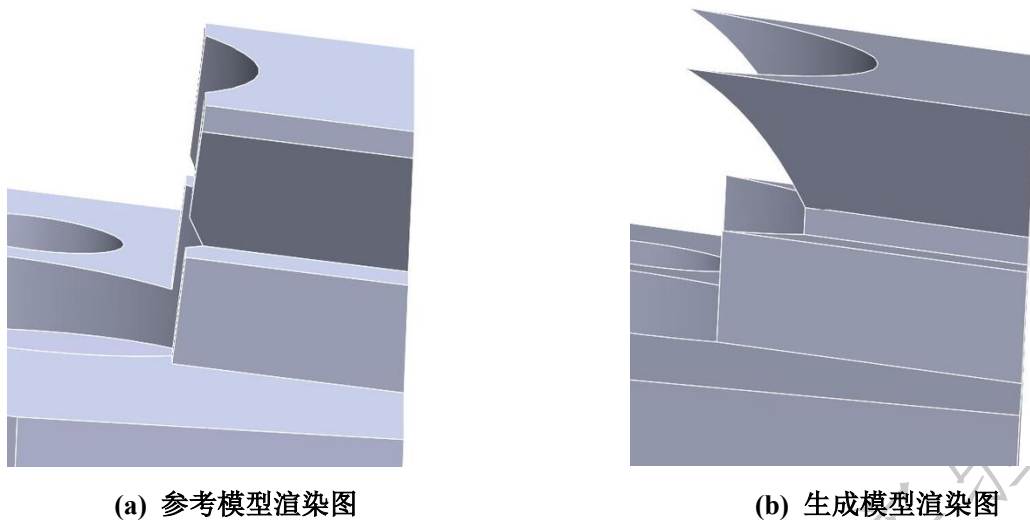


图4-13 D2C-A01工程图生成CAD细节

D2C-A02台钳装配体（复杂）

结果展示

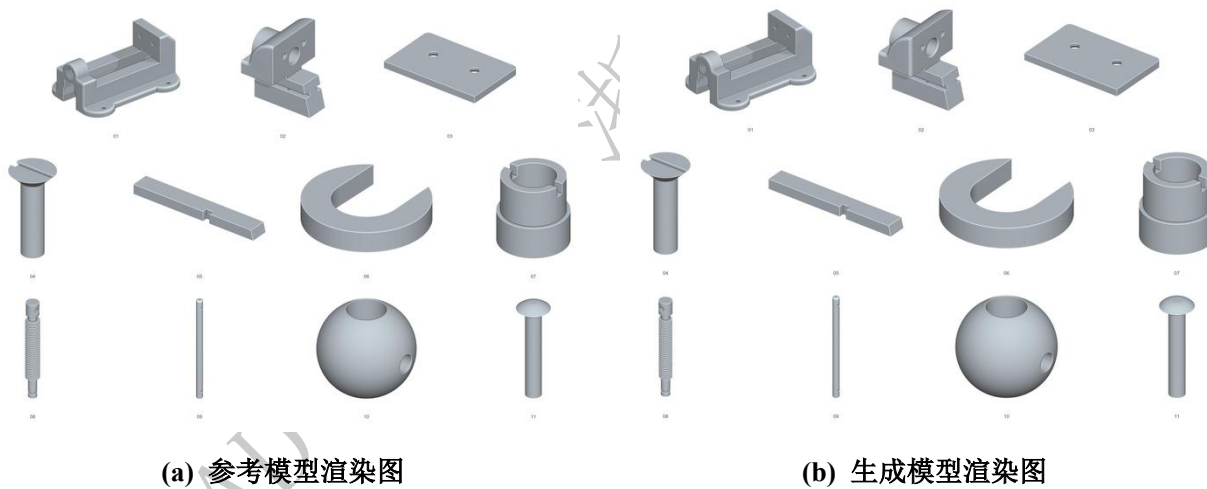


图4-14 D2C-A02工程图生成CAD结果

装配验证

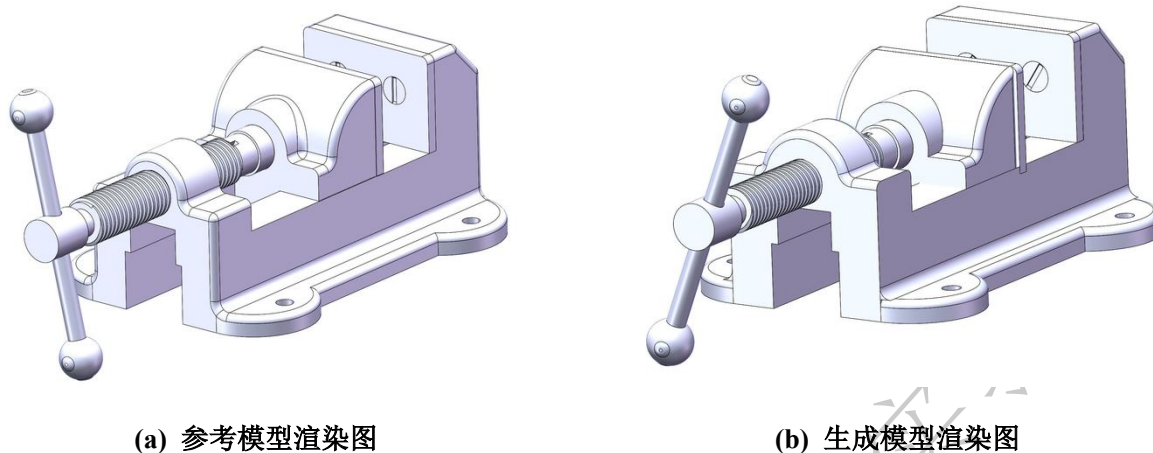


图4-15 D2C-A02工程图生成CAD装配结果

结果分析：生成模型各零件实体有效，其主体结构、主要特征及整体装配关系均实现了较完整的还原，整体形态与参考模型基本一致，说明模型能够较好理解多零件之间的结构对应关系。但部分零件尺寸仍存在偏差，导致装配后局部出现几何重叠与干涉，如图4-16所示，螺纹传动区域及内部配合位置均存在不同程度的实体交叠。因此，该样本结构还原和装配关系理解基本正确，但参数准确性不足，尺寸误差已进一步影响实际装配效果。同时，存在类似上图装配体台体圆角倒角处的“圆角下方是另一个圆角”的非正常结构，这进一步说明GPT-6并没有理解倒角/圆角的意义。

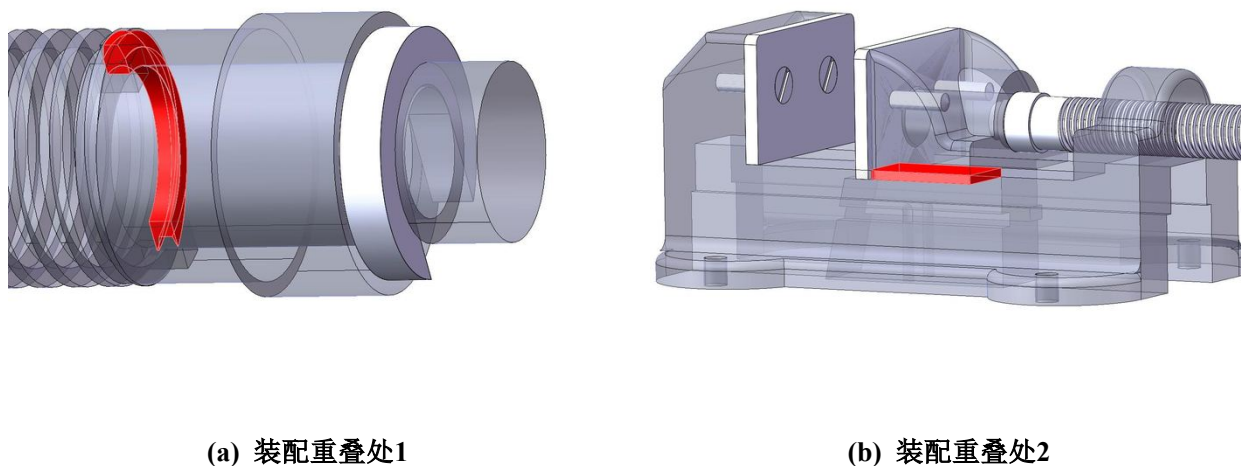


图4-16 D2C-A02工程图生成CAD细节

D2C-A03六缸发动机（复杂）

结果展示



图4-17 D2C-A03工程图生成CAD结果

装配验证 部分零件中用于管路连接及零件对齐的关键孔结构未被正确生成，导致相应管路及连接件无法建立正常装配关系，因此未能完成完整装配验证。

结果分析：生成模型出现三处无效实体，出现非流形边点、不合法的面交叉、甚至存在无实体模型（只有两个曲面）。多数零件的主体外形和基本结构被还原，但部分关键孔系及局部复杂结构存在明显缺失，尺寸参数也存在一定偏差。如图4-18中所示，参考模型中由放样、抽壳等操作形成的复杂曲面、薄壁及过渡结构未得到准确还原，部分区域以较基础的建模形式进行近似替代。尤其是部分承担管路连接和零件定位作用的孔结构缺失，使相关零件失去必要装配接口，最终无法完成完整装配验证；其他地方还存在管道漏孔、错误位置开孔等问题。因此，该样本在整体形态上具备一定还原能力，但复杂结构完整性和参数准确性均未达到要求。

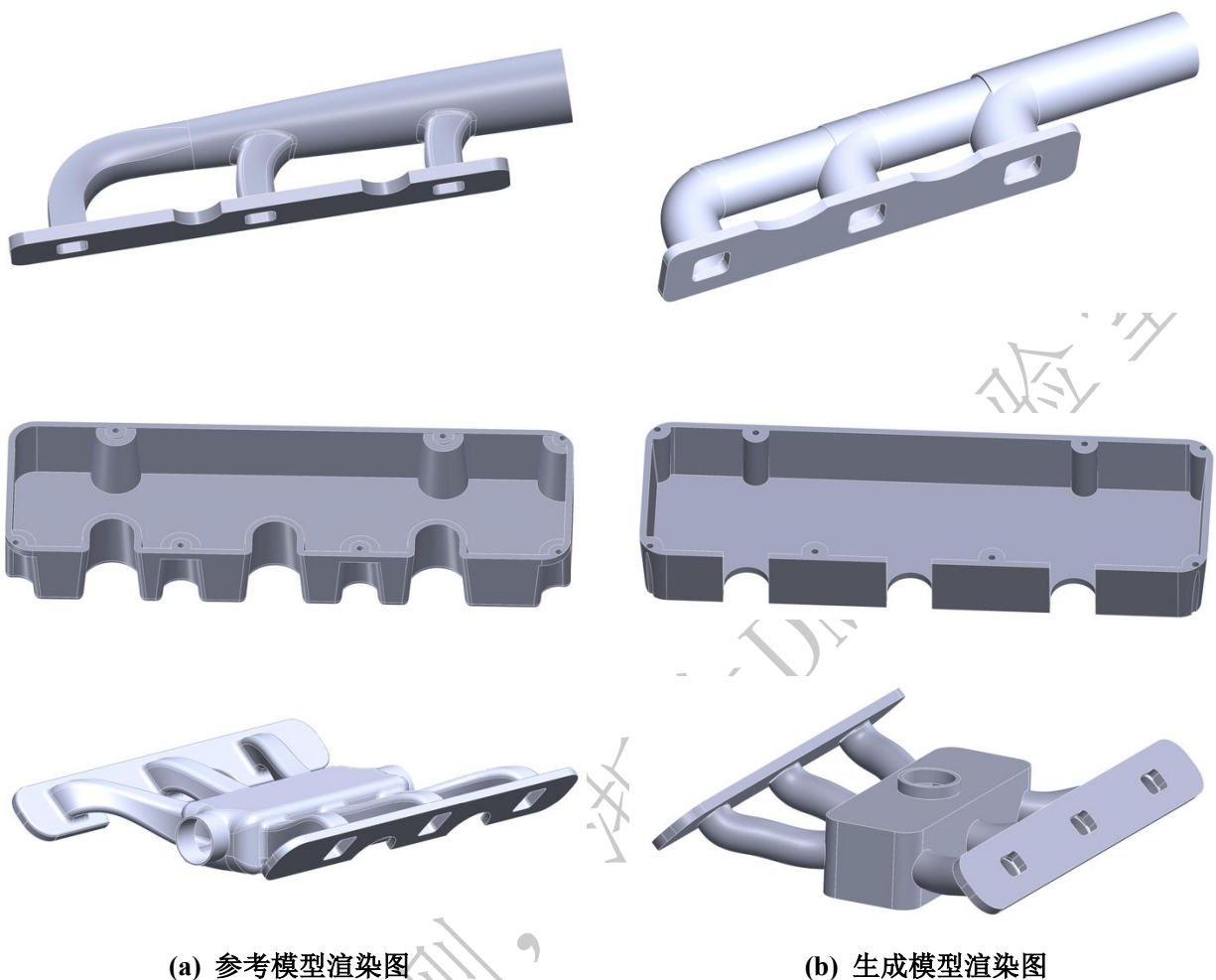


图4-18 D2C-A03工程图生成CAD细节

在测试中还存在一处有趣的发现，输入如4-19所示的工程图时，对比观察8mm半径的杆部可以得知六边形头部厚度应为6mm，但是GPT-6识别时未能注意数字颠倒，以9mm的厚度进行了零件生成。

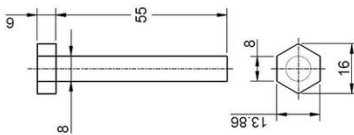


图4-19 D2C-A03被GPT-6识别错误的工程图

4.3.2 扩展实验结果

扩展实验按照4.2.2节设置的尺寸信息缺失和无网络搜索两种条件依次展示。每项实验均以对应样本的主实验结果为对照，仅改变该设置规定的条件，其余输入处理、Prompt及测试环境保持一致。

(1) 尺寸信息缺失

本设置仅保留总长、总宽、总高的包围盒尺寸，删除其余尺寸标注。以下分别展示所选样本在主实验与尺寸信息缺失条件下的输入和结果，并分析局部尺寸缺失带来的变化。各测试样本的用时见表4-8。

表4-7 尺寸信息缺失时工程图到CAD模型重建/生成任务扩展实验结果概览

样本ID	难度	实体有效性	结构还原	尺寸准确	装配效果
D2C-P01	简单	通过	通过	通过	-
D2C-P03	中等	通过	稍有偏差	稍有偏差	-
D2C-A01	中等	通过	三个模型结构错误	出现一定偏差	无法进行装配
D2C-A02	复杂	通过	基本还原	出现一定偏差	装配产生重叠

表4-8 尺寸信息缺失实验各样本测试用时

样本ID	S1_D2C-P01	S1_D2C-P03	S1_D2C-A01	S1_D2C-A02
用时	18分钟	24分钟	19分钟	34分钟

D2C-P01通用支架（简单）

结果对比

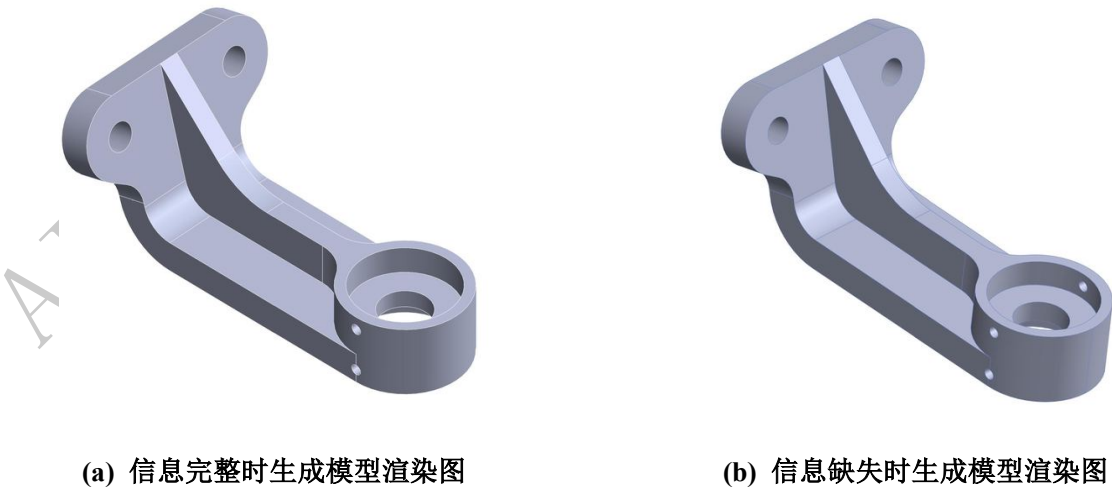


图4-20 D2C-P01尺寸信息缺失工程图生成CAD对比结果

对比分析：信息缺失时生成的模型很好地构建出工程图对应模型，并且与信息完整时的生成相比，之前缺失的孔洞都有生成。

D2C-P03 TITAN-513M（中等）

结果对比

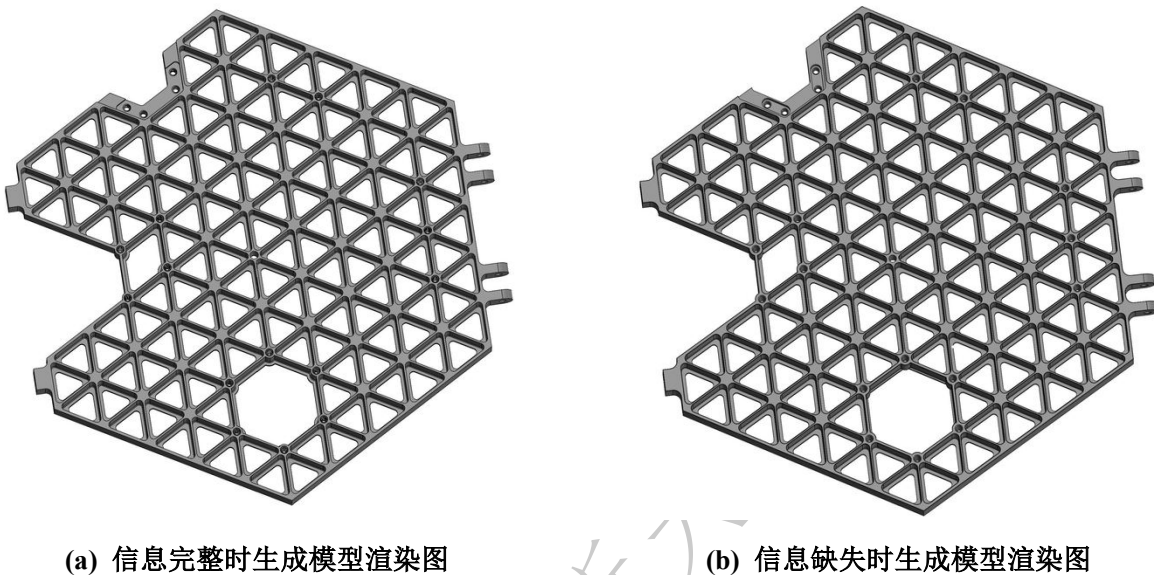
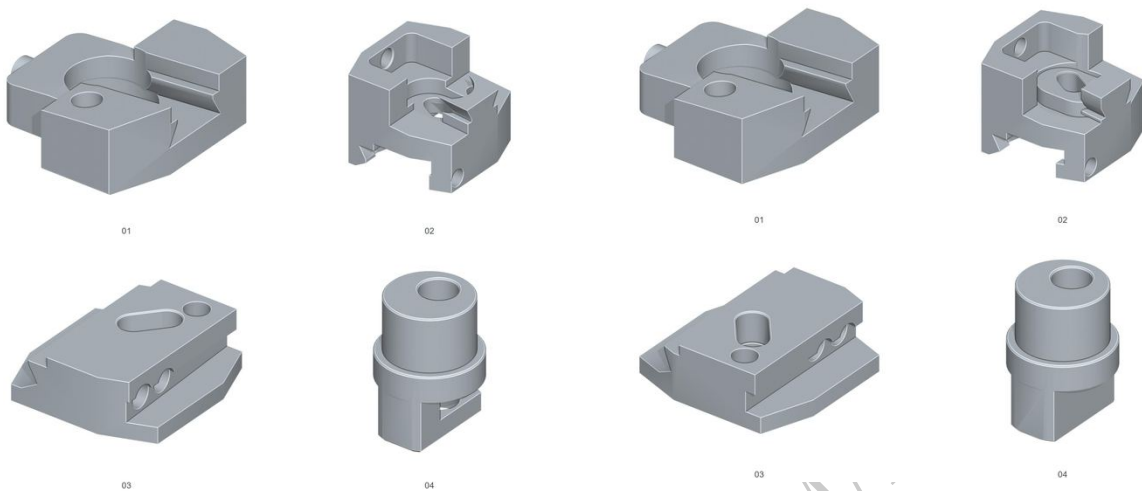


图4-21 D2C-P03尺寸信息缺失工程图生成CAD对比结果

对比分析：信息缺失时的生成模型仍然近乎完整构建出工程图对应模型（螺纹处因删除标注故没有构建，部分厚度偏差在允许范围内）。同时一个有趣的发现是模型对于应当存在螺纹的孔处进行了倒角，说明捕捉到了一定工程图俯视图的细节。

D2C-A01技能五轮铣削滑动装配件（中等）

结果对比



(a) 信息完整时生成模型渲染图

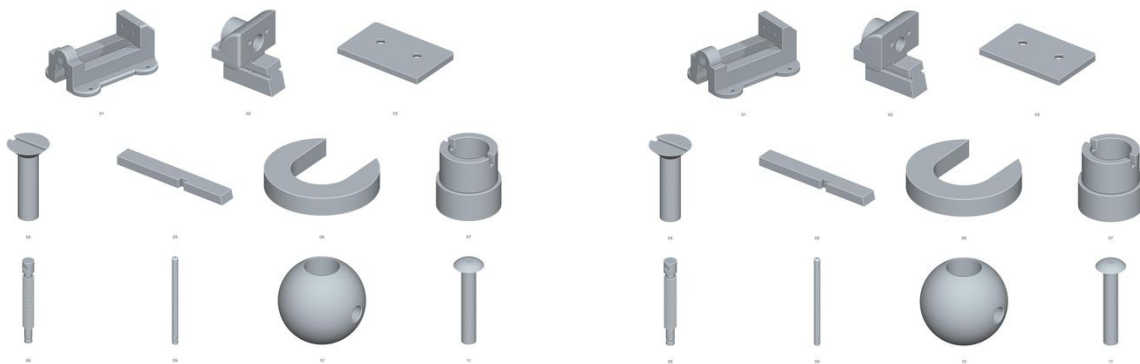
(b) 信息缺失时生成模型渲染图

图4-22 D2C-A01尺寸信息缺失工程图生成CAD对比结果

对比分析：信息缺失时生成的模型能够依据完整视图恢复4个零件的总体轮廓及大部分主要结构，但局部尺寸信息缺失后，特征的具体形状、位置及相互尺寸关系出现了更明显的偏差，例如图中的孔、槽及局部型腔结构与信息完整条件下存在较大差异。

D2C-A02台钳装配体（复杂）

结果对比



(a) 信息完整时生成模型渲染图

(b) 信息缺失时生成模型渲染图

图4-23 D2C-A02尺寸信息缺失工程图生成CAD对比结果

对比分析：在尺寸信息缺失条件下，生成结果仍基本还原了台钳装配体各零件的主体结构与整体外形。然而，与信息完整条件下的生成结果相比，局部尺寸及配合参数出现了

更明显的偏差，导致装配效果显著下降。如图4-24所示，部分配合区域出现了实体重叠、接触面未能正确贴合等问题，反映出尺寸缺失主要影响了零件间的参数匹配与装配精度。

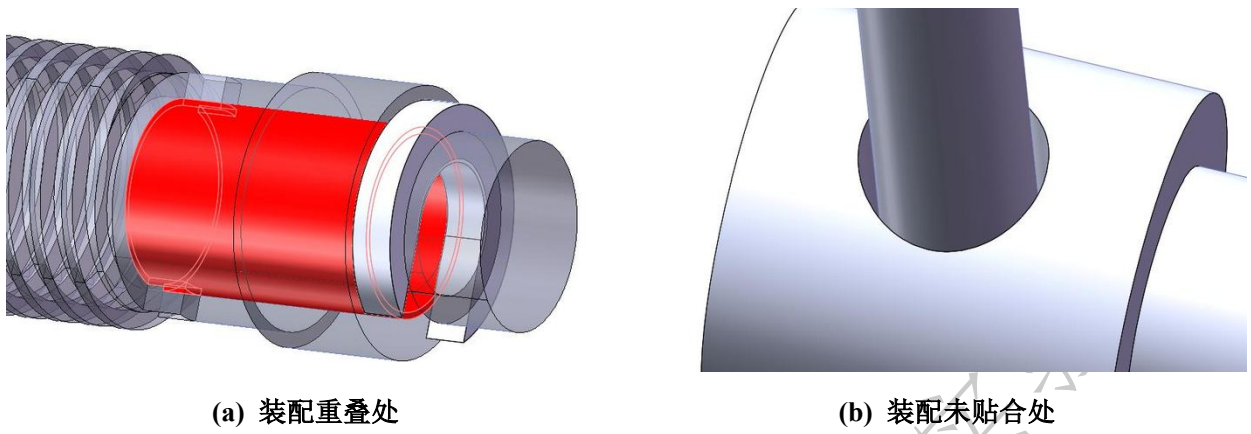


图4-24 D2C-A02尺寸信息缺失工程图生成CAD细节

(2) 无网络搜索

本设置关闭网络搜索权限，输入工程图和Prompt与主实验保持一致。以下直接对比相同样本在联网与无网络条件下的生成结果。各测试样本的用时见表4-10。

表4-9 无网络搜索时工程图到CAD扩展实验结果概览

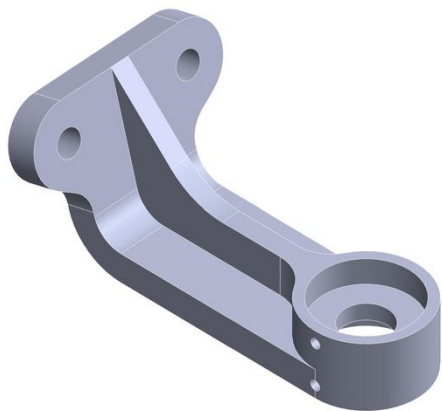
样本ID	难度	实体有效性	结构还原	尺寸准确	装配效果
D2C-P01	简单	通过	通过	通过	-
D2C-P03	中等	通过	稍有偏差	稍有偏差	-
D2C-A01	中等	3/4	基本还原	存在一定偏差	装配存在重叠
D2C-A02	复杂	通过	部分结构缺失	存在一定偏差	无法进行装配

表4-10 无网络搜索实验各样本测试用时

样本ID	S2_D2C-P01	S2_D2C-P03	S2_D2C-A01	S2_D2C-A02
用时	21分钟	46分钟	21分钟	45分钟

D2C-P01通用支架（简单）

结果对比



(a) 允许网络搜索时生成模型渲染图



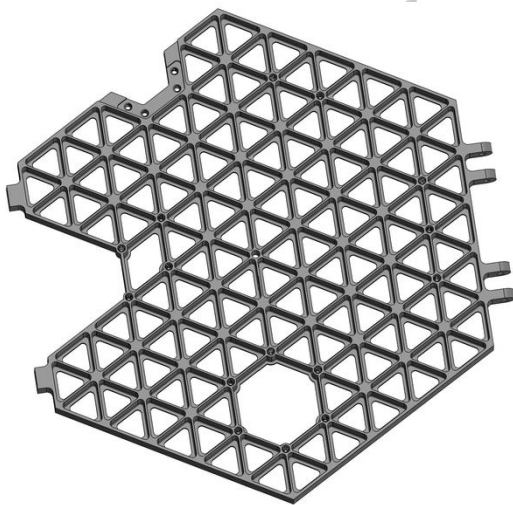
(b) 无网络搜索时生成模型渲染图

图4-25 D2C-P01无网络实验工程图生成CAD对比结果

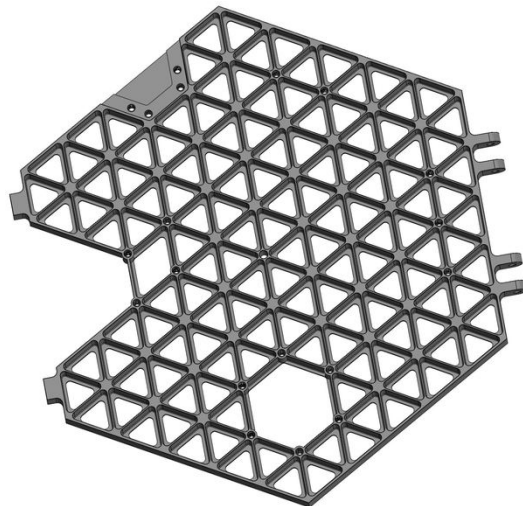
对比分析：无网络搜索时生成的模型很好地构建出工程图对应模型，并且与允许网络搜索时的生成相比，之前缺失的孔洞都有生成。

D2C-P03 TITAN-513M（中等）

结果对比



(a) 允许网络搜索时生成模型渲染图



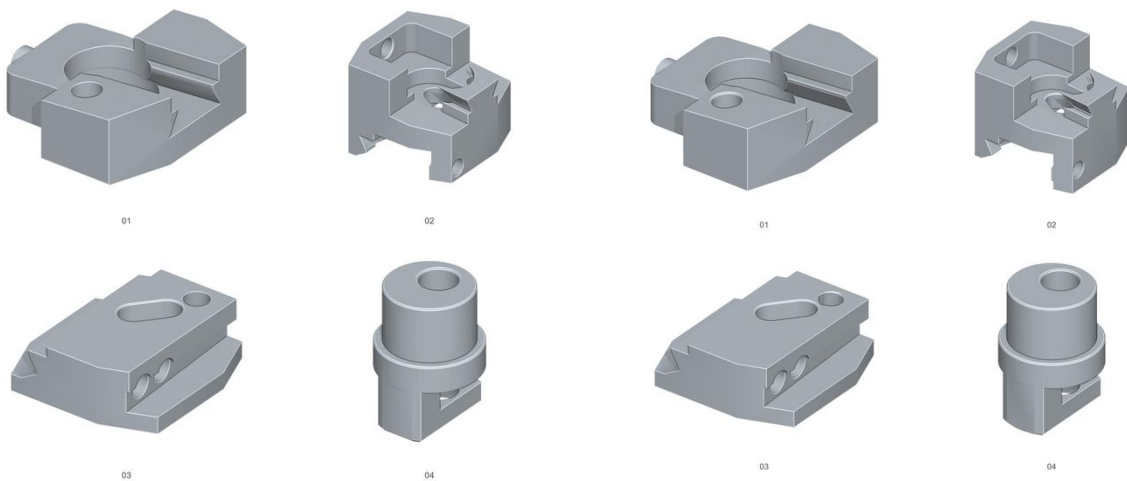
(b) 无网络搜索时生成模型渲染图

图4-26 D2C-P03无网络实验工程图生成CAD对比结果

对比分析：生成模型基本构建出工程图对应模型，但存在一些结构和细节差异。比较惊喜的是之前出现的螺纹孔倒角问题没有出现。

D2C-A01技能五轮铣削滑动装配件（中等）

结果对比



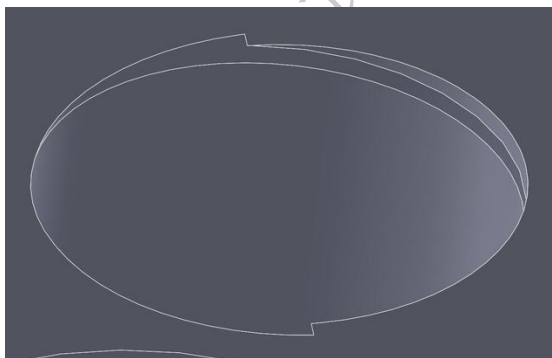
(a) 允许网络搜索时生成模型渲染图

(b) 无网络搜索时生成模型渲染图

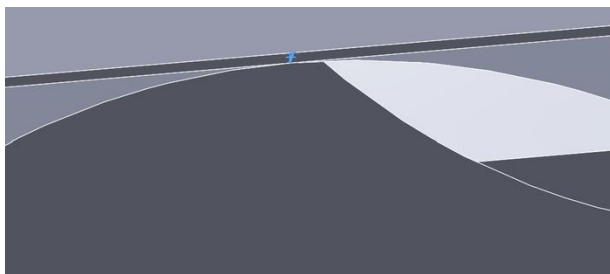
图4-27 D2C-A01无网络实验工程图生成CAD对比结果

对比分析：在无网络搜索条件下，生成模型总体仍能够较好恢复4个零件的主体结构，与允许网络搜索时的结果整体差异较小。但右上角零件的下部结构出现了额外实体，导致该零件局部外形与参考结果不一致，并在后续装配过程中与相邻零件发生几何重叠。

额外需要说明的是，零件上还存在如图4-28(a)所示的某些不自然的孔洞，或是如图4-28(b)所示的线接触导致的非流形边（局部形成非流形连接，不满足本报告采用的实体有效性条件），这是参数与工程图不一致带来的后果。



(a) 不自然形状

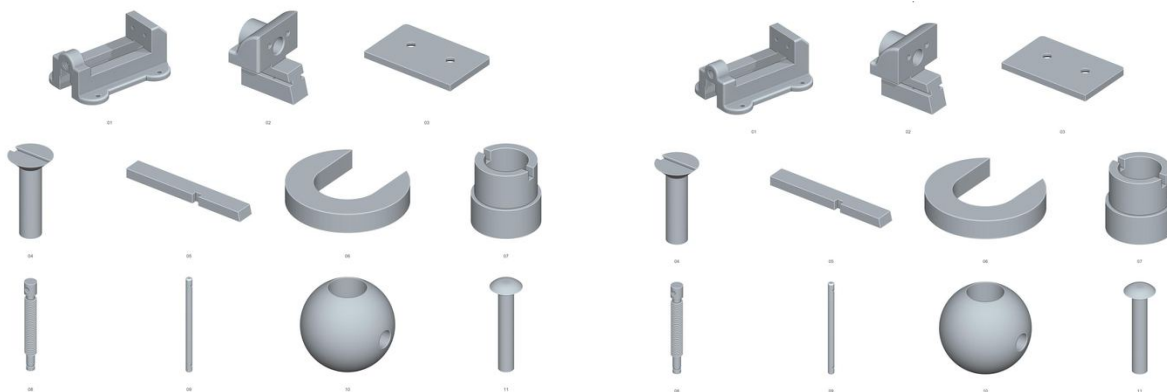


(b) 无效实体（线接触）

图4-28 D2C-A01无网络实验工程图生成CAD细节

D2C-A02台钳装配体（复杂）

结果对比



(a) 允许网络搜索时生成模型渲染图

(b) 无网络搜索时生成模型渲染图

图4-29 D2C-A02无网络实验工程图生成CAD对比结果

对比分析：无网络搜索条件下生成结果与允许网络搜索时整体较为一致，各组成零件的主体结构和主要几何特征均得到较好保留，未出现明显的整体结构错误。主要差异在于部分内螺纹特征未被正确生成，相关孔仅保留了基本孔结构。同时，之前非正确结构的圆角得到了修复。

4.4 GPT-6推理过程分析

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对工程图到CAD模型重建/生成任务可能采用如下推理流程：

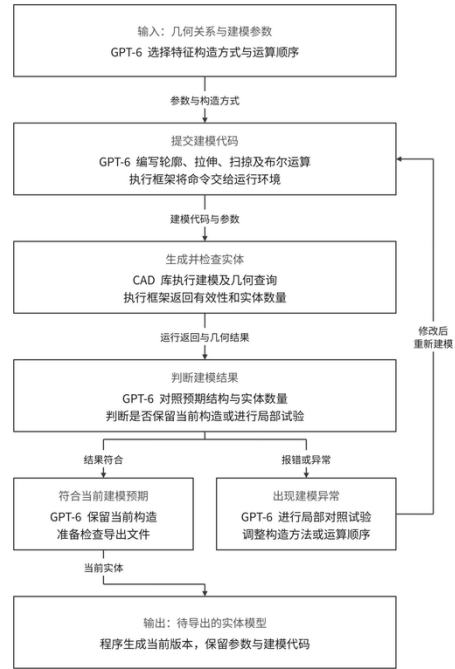
① 工程图与建模参数

确定图纸关系、坐标基准和参数来源



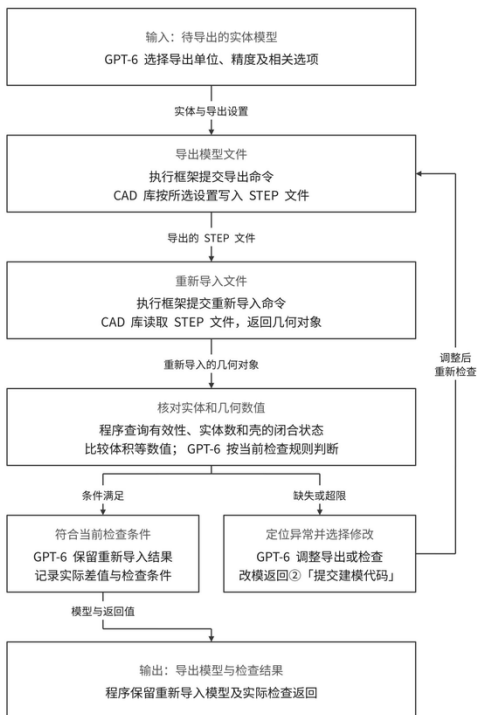
② 实体建模与局部试验

将参数转为实体，并处理建模异常



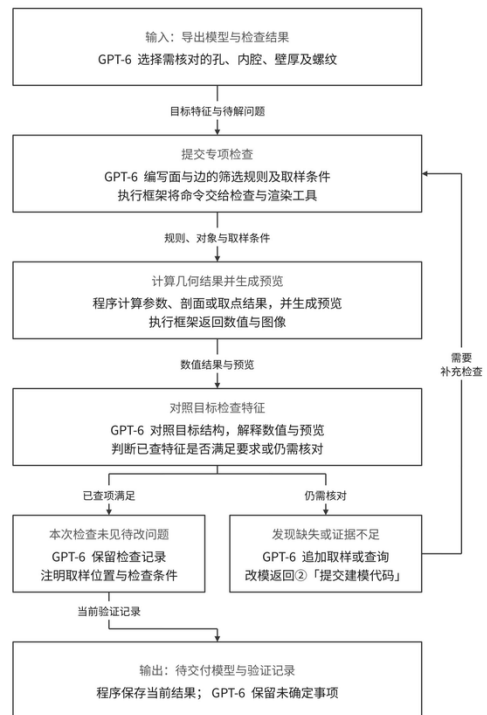
③ 导出与重新导入检查

核对导出文件是否保留实体及几何数值



④ 特征与外观检查

核对关键结构，并据反馈修改模型或补充测量





4.5 能力边界

在工程图到CAD模型重建/生成任务上，GPT-6已能够综合工程图中的多视图投影、剖视信息、尺寸标注及局部特征关系，恢复零件主要三维结构并生成CAD模型。对于简单和中等复杂度样本，主体轮廓、孔槽、重复阵列及大部分内部结构通常能够正确还原；对于发动机缸体等复杂单零件，也能够识别主要孔系、腔体、加强结构及其空间关系。在多零件场景下，GPT-6能够分别完成零件建模，并理解零件之间的对应与装配关系，中等复杂度样本可完成正常装配，复杂装配体的主要组成和整体结构也能够得到较完整恢复。

当前主要能力边界集中在复杂局部结构理解、尺寸参数恢复及装配接口一致性。当工程图中出现多层台阶、型腔嵌套、孔腔组合、通盲孔区分以及倒角、圆角等局部过渡结构时，GPT-6容易出现层次简化、特征遗漏或局部拓扑错误。部分模型虽整体外形基本正确，但具体深度、壁厚、孔位及配合尺寸仍存在偏差；这些误差在单零件中主要表现为参数不准确，在多零件任务中则会进一步累积为装配干涉、无法贴合，甚至因关键孔系缺失而失去必要的装配接口。复杂样本中还出现非法面相交、非流形结构及无有效实体等问题，说明模型尚不能稳定保证复杂CAD几何的合法性，也缺乏对局部错误和装配干涉的可靠自检与修正。

此外，GPT-6对抽壳、拔模、放样等高层CAD建模操作的使用仍不够稳定。在部分复杂样本中，本应通过上述操作形成的薄壁、变截面和复杂过渡结构，被拉伸、切除等基础

几何操作近似替代。此类方法虽可在一定程度上恢复整体外形，但难以准确保持壁厚、曲面过渡及局部接口关系，容易造成局部失真或结构缺失。该现象既可能反映模型对高层CAD特征及其工程含义理解不足，也可能与所采用的建模CAD工具有关。

总体来看，GPT-6已具备较强的工程图多视图理解和三维CAD模型重建能力，但距离稳定的工程级能力仍有明显差距。当前瓶颈已不再是“能否看懂零件大致长什么样”，而是能否准确恢复复杂局部结构、采用合适的三维建模表达、严格落实尺寸与配合关系，并保证生成结果的几何合法性和装配一致性。

AI+CAD评测，

浙大DMA实验室

5 文本到CAD模型生成

5.1 任务目标与评价标准

文本到CAD模型生成包括几何要求建模和功能需求建模两类任务。几何要求建模以目标形状、尺寸、数量和空间关系的文本描述为输入，重点测试GPT-6能否理解这些要求，自主规划建模过程，并调用CAD软件API生成相应的B-rep模型。功能需求建模以使用对象、功能目标、工况、接口和空间限制的文本描述为输入，重点测试GPT-6能否据此选择结构方案、确定设计参数，并完成CAD建模。

主实验在允许网络搜索的条件下开展。扩展实验设置几何要求建模压力测试、功能需求建模压力测试和无网络搜索实验，分别考察GPT-6处理复杂任务以及独立依据给定文本建模的能力。具体任务目标与评价标准见表5-1。

表5-1 文本到CAD模型生成任务目标与评价标准

测试要素	说明
测试问题	GPT-6能否依据几何要求生成形状、尺寸和空间关系一致的CAD模型？能否依据功能需求确定合理的结构和参数，并生成满足设计要求的CAD模型？
核心输入	英文几何要求或功能需求文本。前者描述目标形状、尺寸、数量和空间关系；后者描述使用对象、功能目标、工况、接口和限制条件。
核心输出	根据输入文本生成的B-rep模型，以STEP文件交付。单实体零件应构成有效实体；装配体应保留独立零件及其装配关系。
评价标准	从实体有效性和输入要求满足情况两方面评价生成模型。几何要求的核对内容包括形状、尺寸、数量和空间关系；功能需求的核对内容包括结构、接口、运动空间和设计参数。

5.2 测试用例设计

5.2.1 测试样本选取

两类任务均设置简单、中等和复杂三个档次，每档包含一个单实体零件和一个装配体，共12个主实验样本。本测试依据输入要求所涉及的结构及其设计、建模难度划分样本复杂度。几何要求建模的难点主要在于形状构造和部件之间的空间关系；功能需求建模的难点主要在于结构选择及多项要求的相互制约。

（1）几何要求建模样本

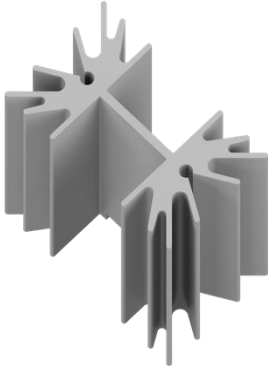
几何要求建模的复杂度取决于形状是否规则、曲面与内部腔道是否易于构造，以及各部件的位置和连接关系是否复杂。单实体零件的复杂度评价侧重局部特征与主体之间的连接和过渡；装配体的复杂度评价侧重零件配合、机构连接和整体布置。标准见表5-2。

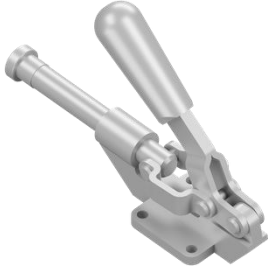
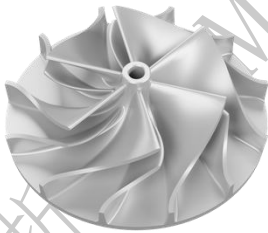
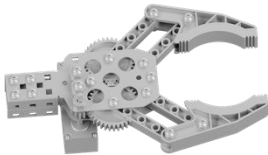
表5-2 几何要求建模样本复杂度划分标准

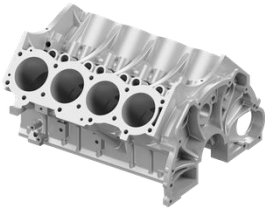
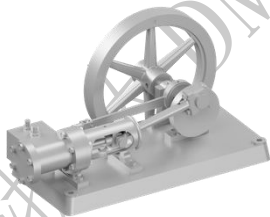
样本复杂度	选取标准	典型表现
简单	形状以等截面、对称或重复结构为主，局部特征之间相互影响较少；装配体主要由一组简单机构组成。	主体形状规则，孔槽较少，结构对称或重复；零件之间的定位和连接关系较直接。
中等	模型含有扭转曲面或连续过渡结构，或同一机构包含多组传动和连接关系。	曲面具有弯曲或扭转变化的，局部连接需要平滑过渡；多个零件通过齿轮、连杆等共同传动。
复杂	内部腔道密集，壁面与孔道的位置相互制约；装配体包含多类部件和多组机构，各部件的空间布置相互影响。	孔道方向多、腔体分布密集，局部壁厚和连接关系复杂；多组机构共用有限空间，位置关系紧密。

几何要求建模选取电子散热器主体、推拉式快速夹具、曲面叶轮、VEX机器人夹爪、Rover V8缸体和卧式单缸蒸汽机六个样本，参考模型来自GrabCAD公开机械模型。测试输入为根据参考模型编写的英文几何要求，描述目标尺寸、形状、数量及相互位置。样本清单见表5-3，表中展示文本输入片段截图和参考模型，并说明各样本的复杂度划分依据。

表5-3 几何要求建模测试样本清单

样本ID	样本类型	文本输入片段截图	参考模型渲染图	复杂度	复杂度归类说明
TC-01	单实体零件	Create a single solid heat-sink body approximately 63.5 mm long, with a 42 × 25.4 mm overall cross-section. A thin, flat central web connects two mirrored groups of longitudinal fins. Each group contains eight fins, arranged as four above and four below the web.		简单	散热片对称重复，主体沿长度方向截面一致，孔槽位置明确。

样本ID	样本类型	文本输入片段截图	参考模型渲染图	复杂度	复杂度归类说明
TC-02	装配体	Create an assembled push-pull toggle clamp approximately 148 mm long, 40 mm wide and 101 mm high. A square mounting plate, approximately 40 × 40 × 5 mm, has rounded corners and four symmetrically arranged mounting holes. A raised longitudinal support carries a cylindrical guide at the front and a fixed hinge at the rear.		简单	夹具由一组推拉机构组成，推杆与导向件需同轴，各铰接轴需平行。
TC-03	单实体零件	Create a single solid open impeller approximately 67.2 mm in overall diameter and 22.9 mm in axial depth. It has a circular backplate approximately 64.8 mm in diameter, a concentric 4.1 mm through-bore and a slender central nose.		中等	两类叶片的曲率、高度和径向范围不同，且需与轮毂、背盘平滑连接。
TC-04	装配体	Create an open two-jaw robotic gripper approximately 195 mm long, 120.5 mm wide and 67.3 mm thick. Two parallel perforated support plates carry a planar gear-and-linkage mechanism, with portions of the gear rims exposed. A perforated mounting tail extends rearward, and a rectangular motor housing sits outside one support plate.		中等	齿轮和多组连杆共同控制夹指位置，啮合、铰接及平行关系相互关联。

样本ID	样本类型	文本输入片段截图	参考模型渲染图	复杂度	复杂度归类说明
TC-05	单实体零件	Create a single solid bare V8 cylinder block approximately 520 mm long, 475 mm wide and 290 mm high. Two cylinder banks form a 90° V, each containing four parallel bores approximately 92 mm in diameter and 110 mm deep. The cylinders repeat at approximately 107.5 mm longitudinal spacing, with the two banks slightly staggered.		复杂	缸孔、水套及其他孔道密集分布，各通道需按要求连通或隔开，孔道壁面需保持连续。
TC-06	装配体	Create a complete horizontal single-cylinder steam-engine assembly on a rectangular base approximately 170 × 80 × 14.5 mm, with rounded corners, mounting holes and a recessed underside. The cylinder lies along the long direction of the base and extends beyond one end.		复杂	主传动与配气机构共用主轴，气缸、连杆和飞轮的位置相互制约。

（2）功能需求建模样本

功能需求建模的复杂度主要依据实现各项功能时需要同时考虑的条件及其相互影响划分。简单任务通常围绕支承或夹持等主要用途展开；中等任务需要在同一机构或部件中兼顾运动、受力、接触或流动要求；复杂任务则涉及多种功能，结构调整往往同时影响空间布置、运动状态和性能。具体划分标准见表5-4。

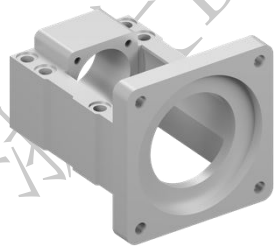
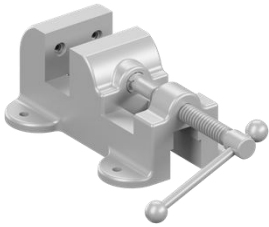
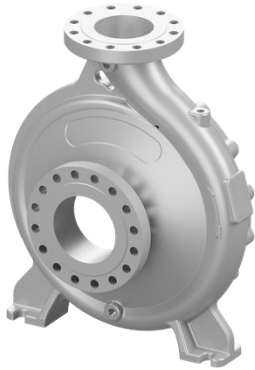
表5-4 功能需求建模样本复杂度划分标准

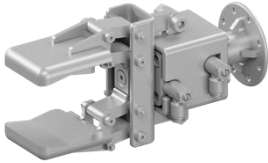
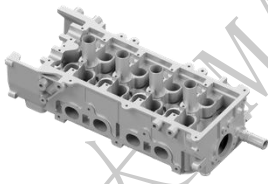
样本复杂度	选取标准	典型表现
简单	主要用途明确，安装、承载和维护要求集中于一个部件或一组简单机构。	以支承、定位或夹紧等用途为主，结构较直接，部件之间的功能关系简单。
中等	同一机构或部件需满足多项相关要求，局部结构的选择会同时影响运动、受力、接触或流动。	导向、调节、传动等结构组合使用，尺寸和配合关系同时影响行程、载荷或接触状态。

样本复杂度	选取标准	典型表现
复杂	多种功能共同受到空间、载荷或工况限制，需要从整体上协调结构和参数。	多个部件或机构协同工作，运动、承载与空间要求相互制约，局部结构变化会影响整机性能。

功能需求建模选取电机与滚珠丝杠同轴支架、手动台虎钳、端吸离心泵蜗壳、气动连杆夹爪、四缸发动机缸盖本体和六足机器人机械底盘六个样本。参考模型均来自GrabCAD，其中支架和缸盖分别采用HM15-86与Nissan CR14DE模型。输入文本给出功能目标、工况、接口和空间限制。样本清单及复杂度划分依据见表5-5。

表5-5 功能需求建模测试样本清单

样本ID	样本类型	文本输入片段截图	参考模型渲染图	复杂度	复杂度归类说明
TRC-01	单实体零件	Design a one-piece motor and ball-screw support for a compact feed mechanism. The motor has an 86 mm square mounting face, a 73 mm diameter locating pilot projecting 2 mm, and four 6.6 mm through-holes on a 69.6 mm square pattern for M6 mounting bolts.		简单	安装接口和刚度要求均围绕电机与丝杠的同轴支承展开。
TRC-02	装配体	Design a manually operated bench vise for holding small steel and aluminium workpieces during filing, sawing, and assembly. The workpieces have opposing flat gripping surfaces and a clamped thickness of 10–60 mm.		简单	台虎钳通过调节、夹紧和释放机构完成工件装夹，功能关系较直接。
TRC-03	单实体零件	Design a one-piece cast-iron casing for a single-stage end-suction pump conveying clean water at 10–60 °C. The application has a nominal pump duty of 500 m³/h at 30 m total head.		中等	泵壳需兼顾流道形状、承压能力及转子安装和拆卸空间。

样本ID	样本类型	文本输入片段截图	参考模型渲染图	复杂度	复杂度归类说明
TRC-04	装配体	Design a pneumatic two-finger gripper for transferring dry aluminium workpieces weighing up to 1 kg. Each workpiece has opposing flat gripping areas of at least 30×30 mm, separated by 20–60 mm, with its centre of mass within 20 mm of the midpoint between the contact areas.		中等	传动机构同时影响开口范围和夹持力，接触垫还需兼顾防滑与表面防护。
TRC-05	单实体零件	Design a one-piece aluminium cylinder-head body for a naturally aspirated, water-cooled, inline-four development engine with a 73 mm bore, 82.8 mm stroke, and 79 mm cylinder spacing.		复杂	缸盖需兼顾承压、散热和配气机构布置，内部通道密集。
TRC-06	装配体	Design a six-legged, battery-powered chassis for indoor inspection on dry, hard floors, using 0.5 as the minimum friction coefficient between the feet and ground.		复杂	腿部运动同时影响通过能力、平台稳定和相机视野，质量与续航要求也相互制约。

5.2.2 测试条件设计

(1) 主实验设计

主实验分别采用几何要求文本和功能需求文本。同类任务使用统一Prompt，各样本在独立会话中测试，并开启网络搜索。

几何要求建模主实验的统一Prompt为：“请根据以下几何要求描述生成对应的CAD模型，准确体现描述中的尺寸、形状和空间关系，保存每次迭代后的模型，并将最终模型导出为STEP文件”。

功能需求建模主实验的统一Prompt为：“请根据以下功能需求描述生成满足需求的CAD模型，保存每次迭代后的模型，并将最终模型导出为STEP文件”。

（2）扩展实验设计

扩展实验包括S1几何要求建模压力测试、S2功能需求建模压力测试和S3无网络搜索实验。S1、S2增加任务复杂度，S3改变网络搜索条件，三者均沿用对应主实验的Prompt。各样本在独立会话中测试并单独记录结果，具体设置见表5-6。

S1几何要求建模压力测试采用七缸星型发动机样本，输入为对应参考模型的几何要求描述文本，涉及尺寸、形状及部件间的空间关系。实验沿用几何要求建模主实验的统一Prompt，并保持网络搜索开启，考察GPT-6处理多组机构和复杂空间关系的能力。

S2功能需求建模压力测试采用HK-450电动直升机机械总成样本，输入为对应参考模型的功能需求描述文本，涉及使用目标、工况和性能要求。实验沿用功能需求建模主实验的统一Prompt，并保持网络搜索开启，考察GPT-6在多项功能相互制约时确定结构方案和完成建模的能力。两个样本的文本输入片段截图与参考模型渲染图见图5-1。

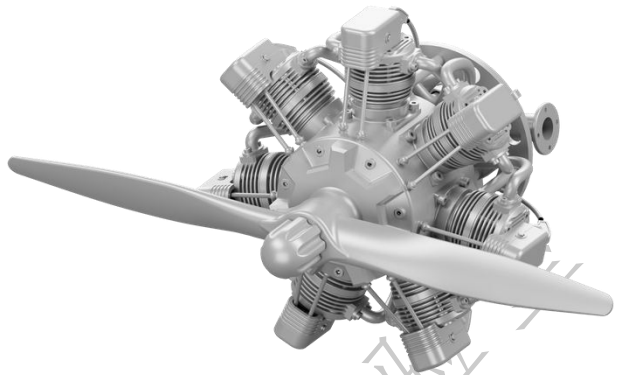
S3无网络搜索实验选取几何要求建模的TC-05、TC-06和功能需求建模的TRC-05、TRC-06，保持样本文本与主实验一致，沿用对应任务的Prompt并关闭网络搜索权限。该实验在Prompt中补充：“本次任务禁止网络搜索及下载外部模型或参考资料，仅依据给定文本和本地工具完成建模”。

表5-6 文本到CAD模型生成扩展实验设计

设置	Prompt	网络搜索	测试目的
S1几何要求建模 压力测试	沿用几何要求建模主实验的统一Prompt。	开启	考察GPT-6能否在多组机构和复杂空间关系下，生成几何特征与装配关系符合要求的模型。
S2功能需求建模 压力测试	沿用功能需求建模主实验的统一Prompt。	开启	考察多项功能相互制约时，GPT-6确定结构方案并生成CAD模型的能力。
S3 无网络搜索	沿用主实验对应任务的Prompt，补充禁止联网、搜索及下载外部资料的指令。	关闭	考察关闭网络搜索后，GPT-6依据给定文本完成复杂模型的能力。

Create a complete single-row, seven-cylinder radial-engine assembly approximately 280 mm long, with a 440 mm propeller span. The cylinder assemblies, including rocker covers, have an overall envelope diameter of approximately 240 mm. Seven cylinders surround the main shaft at equal angles in a common transverse plane, with one upright. The central crankcase is approximately 100 mm in diameter and 67 mm long. Each cylinder has a 24 mm bore, a separate liner, closely spaced annular cooling fins up to approximately 46 mm in diameter, a finned head and a rounded rectangular rocker cover.

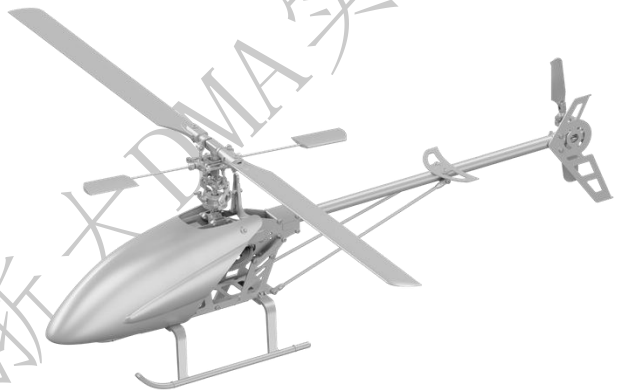
(a) S1文本输入片段截图



(b) S1参考模型渲染图

Design a compact electric helicopter for visual inspection with a 150 g camera module measuring $60 \times 50 \times 40$ mm. The module has four M3 attachment points on a 40×30 mm rectangular pattern and a centre of mass 20 mm above its mounting face. The complete aircraft, including the camera, battery, and electronics, shall weigh no more than 1.4 kg. Its operating envelope, including all rotating and moving components, shall remain within 1,100 mm in length, 850 mm in width, and 350 mm in height.

(c) S2文本输入片段截图



(d) S2参考模型渲染图

图5-1 两类压力测试样本的文本输入片段截图与参考模型渲染图

5.3 测试结果展示

本节先展示几何要求建模和功能需求建模的主实验结果，再展示压力测试和无网络搜索的扩展实验结果。各项实验均评价模型对文本要求的满足情况及实体有效性。

几何要求满足率依据模型的形状、尺寸、数量和空间关系逐项统计；功能需求满足率依据结构、接口、运动空间及相关设计计算逐项统计，反映生成模型及配套设计对输入需求的满足程度。两类任务的判定内容和通过条件分别见表 5-7 和表 5-8。

表 5-7 几何要求满足情况判定标准

判定内容	通过条件
几何特征与部件数量	输入规定的孔、槽、叶片、加强肋等特征及装配部件完整，数量符合要求。
尺寸、角度与位置关系	CAD 模型的外形尺寸、孔径、间距、高度及角度符合输入要求；近似尺寸采用对应样本实际使用的判定容差。
对称、平行与同轴关系	对应特征或部件的布置符合要求，轴线位置和方向满足指定的同轴、平行或对称关系。
曲面形状与局部过渡	受检部位的高度、厚度、后掠及倾斜变化符合文本描述；局部过渡依据实际采用的轮廓、截面或相邻面法向夹角检查结果判定。
截面形状与变化	等截面主体的对应截面在对齐后保持一致，截面差异满足所采用的阈值；变截面结构的轮廓和尺寸变化符合输入要求。
孔道连通与隔离	应贯通的孔道保持连通；应隔离的孔道、腔体之间具有完整隔壁。
装配组成与布局	零件按要求独立建模，安装位置、铰接轴、连接销及指定装配姿态符合文本描述。

表 5-8 功能需求满足情况判定标准

判定内容	通过条件
外形尺寸与安装接口	外形尺寸不超过规定范围；定位孔、安装孔、轴承座等接口的尺寸和位置与给定部件匹配。
部件容纳空间与装配连接	指定部件在检查的安装位置无非预期体积干涉；承担固定或运动传递作用的部件具有完整连接。
运动范围与姿态	开口、行程、转角、计算速度及运动外廓达到输入要求；已检查的运动位置不存在非预期干涉。
支撑稳定性	在检查的支撑姿态和载荷条件下，重心投影位于支撑区域内，计算所需的关节力矩不超过所采用的驱动能力。
受载位移、应力与紧固力	按实际采用的有限元或简化力学模型，受检部位的计算位移、应力不超过对应限值，紧固力能够抵抗规定载荷。

判定内容	通过条件
夹紧力、防滑系数与自锁	规定手力或供气压力下，计算夹紧力达到要求；防滑安全系数达到输入规定的下限；要求自锁时，反向驱动仍需施加正力矩。
接触垫与表面防护结构	软垫承担与工件的接触；在设计采用的压缩量下，金属背板及紧固件仍与工件保留间隙。
流道连通与水头损失	流道连接指定入口和出口；在规定流量及计算采用的阻力系数下，吸入流道总水头损失不超过 0.5 m。
燃烧室容积与压缩比	根据模型腔体容积、给定缸径、行程和垫片尺寸计算，几何压缩比达到 9.9：1。
冷却、润滑结构与热量平衡	水套覆盖规定受热部位，冷却和润滑通道保持连通，燃气、冷却液与机油通道相互隔离；供给接口符合要求，并按给定流量、入口温度和热负荷核算平均温升。
质量、功耗与续航	质量统计不超过规定上限；按采用的可用电量和功耗计算，续航时间及剩余电量达到输入要求。
相机视场与转动范围	已检查姿态下，要求的视场不被模型遮挡；相机机构的转动范围覆盖规定角度，安装及传动连接完整。
调整与拆卸空间	操作工具及待拆部件在已检查的进入、操作和拆卸位置，与保留部件无干涉。

5.3.1 主实验结果

（1）总体概览

两项主实验各包含6个样本，简单、中等和复杂样本各2个，共12个。GPT-6在开启网络搜索的条件下完成了全部样本的STEP模型交付，结果见表5-9和表5-10。对于装配体，评测方将其拆分为单零件STEP后逐一检测实体有效性，表中分数表示通过检测的实体数与实体总数。各测试样本的用时见表5-11。

几何要求建模中，6个样本均满足全部明确列出的几何要求，并通过实体有效性检测。GPT-6能够建立散热片、夹具、叶轮、夹爪、缸体及蒸汽机的主要形状和装配关系，但部分局部连接仍不完整：TC-03的主叶片内缘上段与轮毂之间留有间隙；TC-04的小齿轮与电机输出轴之间缺少传递转矩的固定结构。

功能需求建模中，GPT-6完成了6个样本的结构与参数设计，最终功能需求满足率均为100%。简单样本满足同轴支承、安装维护、工件夹持和自锁要求；中等样本完成了泵壳流道、转子拆卸空间及夹爪传动机构的设计；复杂样本完成了缸盖配气、冷却和润滑结构，以及六足底盘运动、承载和续航的设计。6个样本中有5个通过实体有效性检测。TRC-06的3个电机支承件因实体边界自交未通过检测，模型另含36个未封闭的独立壳体，因此整机未通过实体有效性检测。

表5-9 几何要求建模主实验结果汇总

样本ID	复杂度	实体有效性	几何要求满足情况	测试结果
TC-01	简单	通过	12项全部满足	散热片形状、排列及孔槽位置符合要求
TC-02	简单	通过	10项全部满足	夹具结构完整，推杆、连杆和手柄的位置关系符合要求
TC-03	中等	通过	10项全部满足	叶片数量及排列正确，主叶片内缘上段与轮毂存在间隙
TC-04	中等	通过	12项全部满足	夹指及齿轮布局符合要求，但电机轴与小齿轮缺少传递转矩的固定结构
TC-05	复杂	通过	12项全部满足	缸体结构完整，孔系和内部通道布局符合要求
TC-06	复杂	通过	12项全部满足	整体装配关系正确，主传动与配气机构连接完整

表5-10 功能需求建模主实验结果汇总

样本ID	复杂度	实体有效性	功能需求满足情况	测试结果
TRC-01	简单	通过	8项全部满足	支架满足电机与丝杠的同轴支承及安装维护需求
TRC-02	简单	通过	9项全部满足	台虎钳满足开口调节与工件夹持需求，并能保持夹紧状态

样本ID	复杂度	实体有效性	功能需求满足情况	测试结果
TRC-03	中等	通过	11项全部满足	泵壳满足流道连通、管路连接及转子从后方拆出的要求
TRC-04	中等	通过	9项全部满足	夹爪满足同步开合、工件夹持、表面防护和安装连接要求
TRC-05	复杂	通过	12项全部满足	进排气、冷却和润滑结构完整，配气部件具有运动和检修空间
TRC-06	复杂	未通过 904/907	12项全部满足	行走、转向和越障的运动关系正确；3个电机支承件存在实体边界自交，另有36个开放壳

表5-11 文本到CAD主实验各样本测试用时

样本ID	TC-01	TC-02	TC-03	TC-04	TC-05	TC-06	TRC-01	TRC-02	TRC-03	TRC-04	TRC-05	TRC-06
用时	23分钟	26分钟	35分钟	37分钟	46分钟	47分钟	39分钟	1时20分钟	1时9分钟	1时40分钟	1时40分钟	3时17分钟

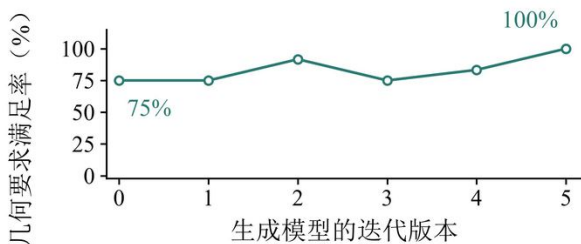
（2）用例分析

以下按任务类型展示各样本的文本输入片段、参考模型、生成模型及要求满足率曲线，并分析生成结果。

① 几何要求建模

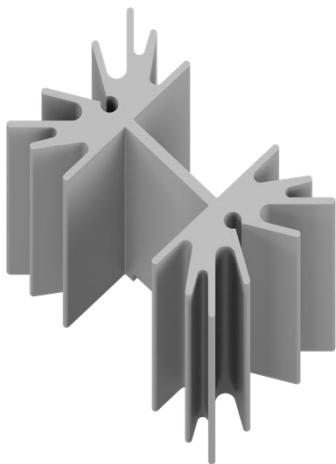
TC-01（简单）

Create a single solid heat-sink body approximately 63.5 mm long, with a 42×25.4 mm overall cross-section. A thin, flat central web connects two mirrored groups of longitudinal fins. Each group contains eight fins, arranged as four above and four below the web. The fins spread outward in a fan-shaped profile, with progressively changing inclinations.

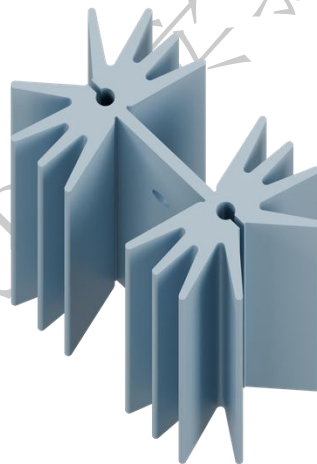


(a) 文本输入片段截图

(b) 几何要求满足率



(c) 参考CAD模型渲染图



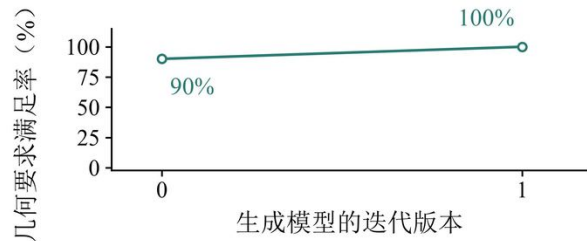
(d) 生成的CAD模型渲染图

图5-2 TC-01几何要求建模结果

结果分析：TC-01要求GPT-6按照给定尺寸和对称关系，生成带扇形散热片、贯通槽和安装孔的单实体散热器。GPT-6在迭代中调整了散热片截面，完善根部过渡，并补齐安装孔和端部缺口。最终模型的16片散热片排列规则，薄片末端与加宽根部之间过渡平滑，两侧C形槽贯穿主体，几何要求满足率由75%提高至100%，见图5-2。最终模型通过实体有效性检测。

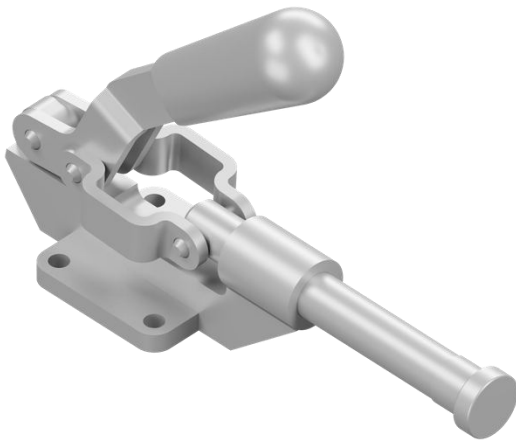
TC-02（简单）

Create an assembled push-pull toggle clamp approximately 148 mm long, 40 mm wide and 101 mm high. A square mounting plate, approximately $40 \times 40 \times 5$ mm, has rounded corners and four symmetrically arranged mounting holes. A raised longitudinal support carries a cylindrical guide at the front and a fixed hinge at the rear.

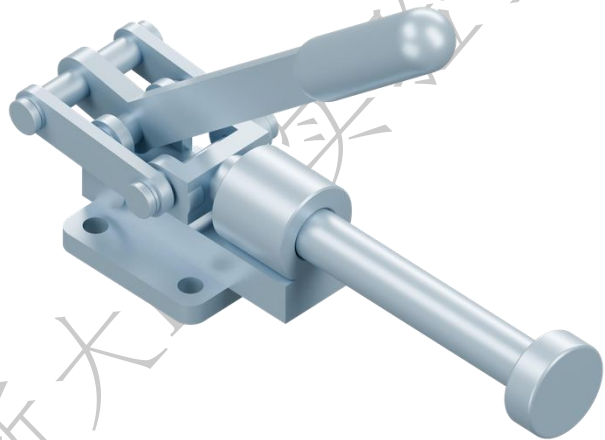


(a) 文本输入片段截图

(b) 几何要求满足率



(c) 参考CAD模型渲染图



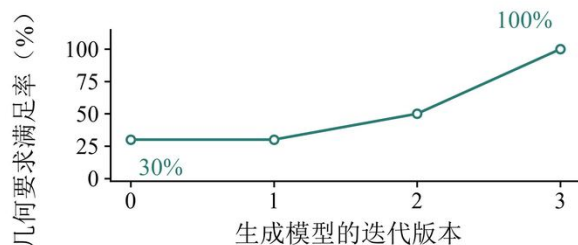
(d) 生成的CAD模型渲染图

图5-3 TC-02几何要求建模结果

结果分析：TC-02要求GPT-6按给定尺寸和装配姿态构建推拉式快速夹具。GPT-6将水平推杆安装在同轴导向件内，通过双侧连杆将推杆与上扬的手柄连接，并在迭代中补齐铰接销和紧固件。最终模型中，连杆接近水平，中间铰接点略高于两端铰接点的连线，符合给定的机构姿态。几何要求满足率由90%提高至100%，见图5-3。夹具的22个零件均通过实体有效性检测。

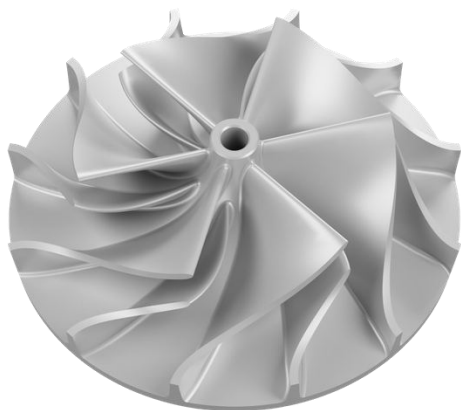
TC-03（中等）

Create a single solid open impeller approximately 67.2 mm in overall diameter and 22.9 mm in axial depth. It has a circular backplate approximately 64.8 mm in diameter, a concentric 4.1 mm through-bore and a slender central nose. The hub rises steeply near the axis and expands smoothly into a broad, shallow profile towards the backplate rim. Six tall main blades alternate with six shorter splitter blades. Six tall main blades alternate with six shorter splitter blades.

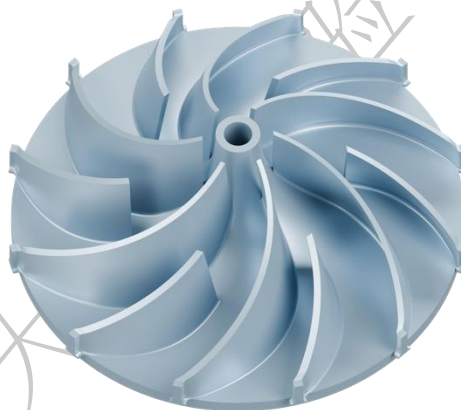


(a) 文本输入片段截图

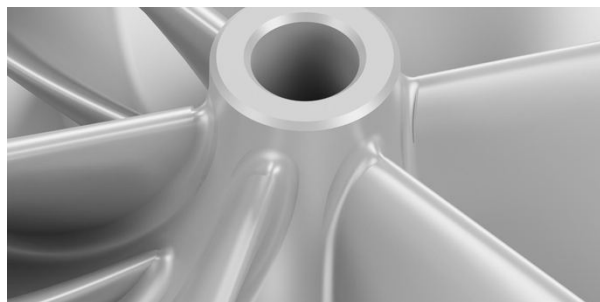
(b) 几何要求满足率



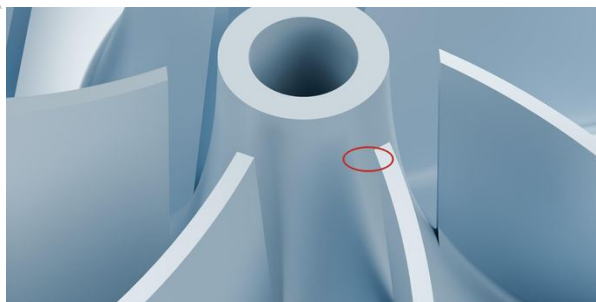
(c) 参考CAD模型渲染图



(d) 生成的CAD模型渲染图



(e) 参考模型：内缘连续连接



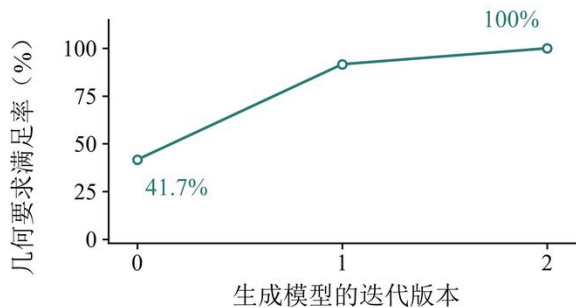
(f) 生成模型：内缘上段留有间隙

图5-4 TC-03几何要求建模结果

结果分析：TC-03要求GPT-6生成具有交替主叶片和分流叶片的单实体叶轮，并满足给定尺寸和扭转形状。GPT-6逐步补齐两类叶片，几何要求满足率由30%提高至100%。局部检查发现，生成模型的叶片下部已与轮毂和背盘连接，但主叶片内缘上段与中央轮毂之间仍留有间隙；参考模型在该处连续连接，见图5-4。GPT-6建立了叶片的主要形状，却没有完整处理叶片与轮毂之间的过渡。最终叶轮通过实体有效性检测。

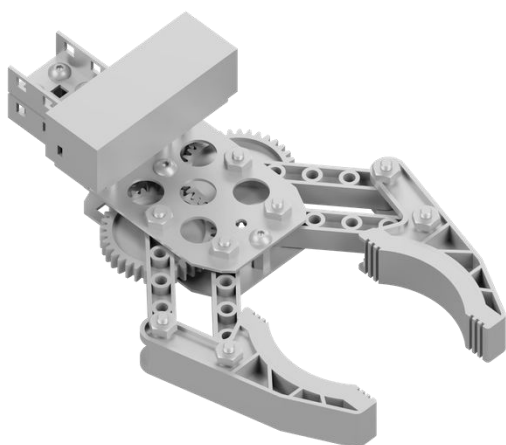
TC-04（中等）

Create an open two-jaw robotic gripper approximately 195 mm long, 120.5 mm wide and 67.3 mm thick. Two parallel perforated support plates carry a planar gear-and-linkage mechanism, with portions of the gear rims exposed. A perforated mounting tail extends rearward, and a rectangular motor housing sits outside one support plate. Two equal intermeshing spur gears have approximately 40.2 mm outside diameters and 38.1 mm centre spacing.

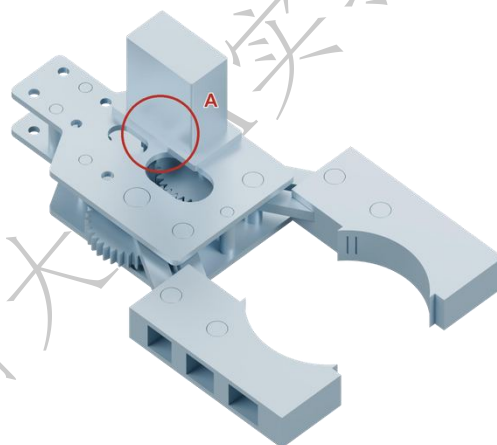


(a) 文本输入片段截图

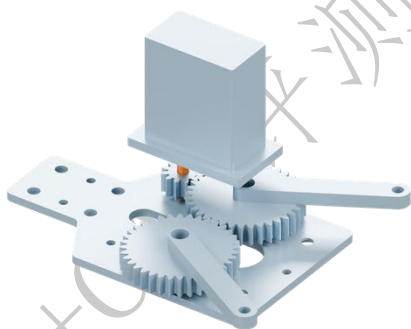
(b) 几何要求满足率



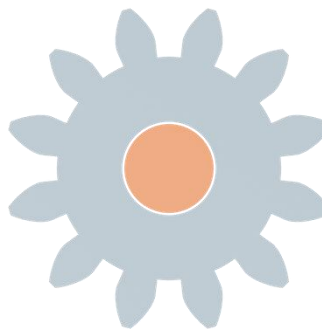
(c) 参考CAD模型渲染图



(d) 生成模型及驱动部位A



(e) A处内部结构（隐藏遮挡支板）



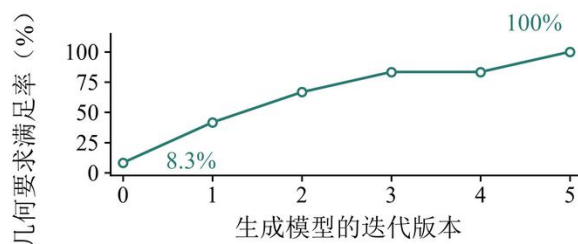
(f) 小齿轮剖面：橙色为电机轴

图5-5 TC-04几何要求建模结果

结果分析：TC-04要求GPT-6建立由齿轮和连杆驱动的双指夹爪，并满足给定尺寸、啮合关系和夹指姿态。GPT-6在迭代中补齐夹指、连杆及安装部件，几何要求满足率由41.7%提高至100%。最终模型的两侧夹指保持平行，但电机与齿轮之间的驱动连接不完整：电机轴穿入小齿轮中心孔后仍留有0.1 mm径向间隙，且没有紧定螺钉等固定结构，小齿轮不能随电机轴可靠转动。图中A处标出了该部位，局部图展示了生成模型内部的实际连接，见图5-5。夹爪的47个零件均通过实体有效性检测。

TC-05（复杂）

Create a single solid bare V8 cylinder block approximately 520 mm long, 475 mm wide and 290 mm high. Two cylinder banks form a 90° V, each containing four parallel bores approximately 92 mm in diameter and 110 mm deep. The cylinders repeat at approximately 107.5 mm longitudinal spacing, with the two banks slightly staggered.

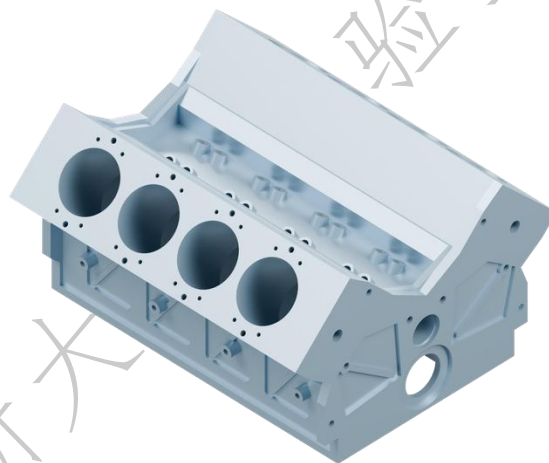


(a) 文本输入片段截图

(b) 几何要求满足率



(c) 参考CAD模型渲染图



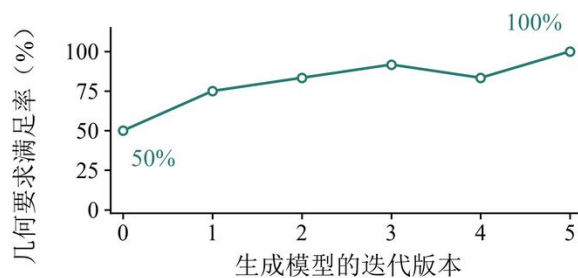
(d) 生成的CAD模型渲染图

图5-6 TC-05几何要求建模结果

结果分析：TC-05要求GPT-6生成包含气缸、水套、轴承座和内部孔道的单实体V8缸体。GPT-6先建立两列气缸和曲轴腔，再构建水套与纵向孔道，并在迭代中修正水套贯通中央凹腔、纵向孔道被封闭的问题。最终模型中，两列气缸呈90°布置，水套连通且与缸孔隔离，底部五道腹板共同支承同轴的曲轴轴承座。几何要求满足率由8.3%提高至100%，见图5-6。最终缸体为单一实体，通过实体有效性检测。

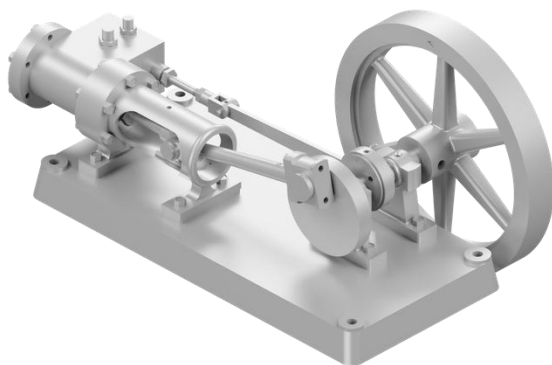
TC-06（复杂）

Create a complete horizontal single-cylinder steam-engine assembly on a rectangular base approximately $170 \times 80 \times 14.5$ mm, with rounded corners, mounting holes and a recessed underside. The cylinder lies along the long direction of the base and extends beyond one end. Its body is 35 mm long, with a 19 mm bore, a narrower barrel approximately 22 mm in diameter and wider end flanges approximately 32 mm in diameter.

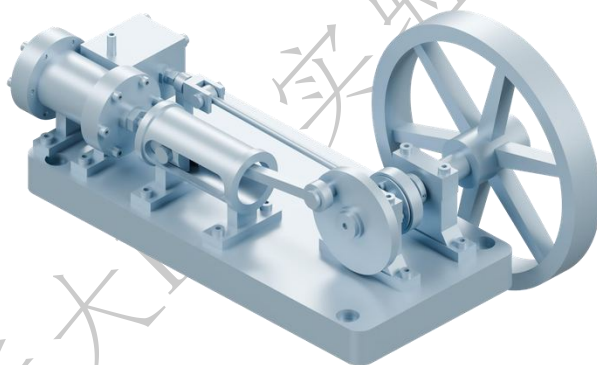


(a) 文本输入片段截图

(b) 几何要求满足率



(c) 参考CAD模型渲染图



(d) 生成的CAD模型渲染图

图5-7 TC-06几何要求建模结果

结果分析：TC-06要求GPT-6按给定尺寸构建卧式单缸蒸汽机，并完成主传动与配气机构的装配。GPT-6逐步建立气缸、曲柄连杆、飞轮和偏心配气机构，补齐紧固件并调整部件位置。最终模型中，活塞杆与十字头沿气缸轴线布置，主连杆连接曲柄，独立的偏心连杆驱动配气机构，各机构的相对位置符合要求。几何要求满足率由50%提高至100%，见图5-7。蒸汽机的106个零件均通过实体有效性检测。

② 功能需求建模

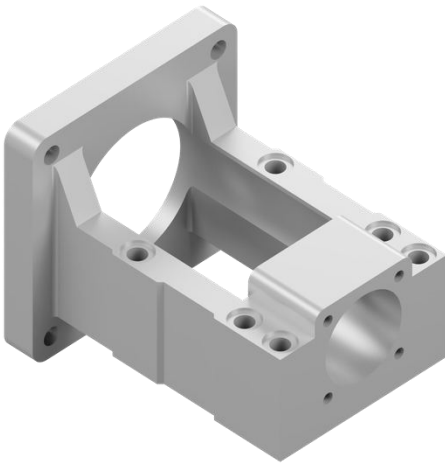
TRC-01（简单）

Design a one-piece motor and ball-screw support for a compact feed mechanism. The motor has an 86 mm square mounting face, a 73 mm diameter locating pilot projecting 2 mm, and four 6.6 mm through-holes on a 69.6 mm square pattern for M6 mounting bolts. The screw support uses a matched pair of angular-contact bearings, each with a 15 mm bore, 32 mm outside diameter, and 9 mm width.

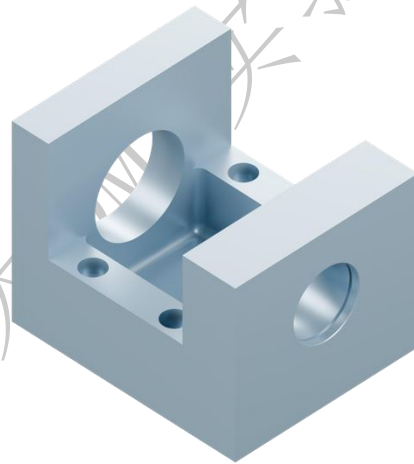


(a) 文本输入片段截图

(b) 功能需求满足率



(c) 参考CAD模型渲染图



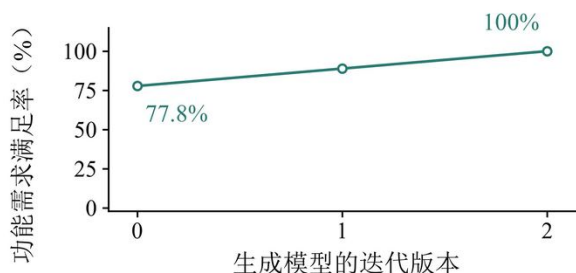
(d) 生成的CAD模型渲染图

图5-8 TRC-01功能需求建模结果

结果分析：TRC-01要求GPT-6设计兼顾同轴定位、承载刚度和装调空间的一体式电机与丝杠支架。GPT-6采用两端支承、中部敞开的结构，使电机定位孔与轴承座保持同轴，并为联轴器调节和轴承更换留出通道。在规定载荷下，计算得到的两轴相对径向偏移约为0.0033 mm，小于0.05 mm限值。首次保存的模型已满足8项功能设计要求，后续迭代的满足率保持100%，见图5-8。最终支架为单一实体，通过实体有效性检测。

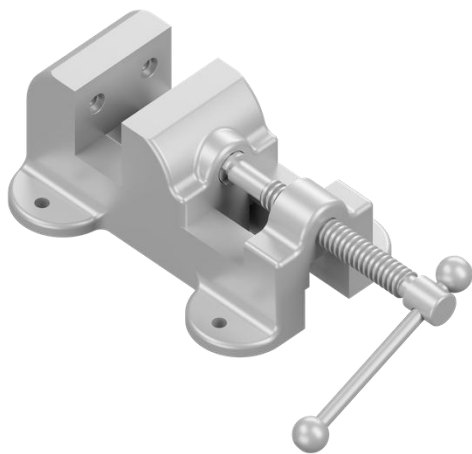
TRC-02（简单）

Design a manually operated bench vise for holding small steel and aluminium workpieces during filing, sawing, and assembly. The workpieces have opposing flat gripping surfaces and a clamped thickness of 10–60 mm. Provide at least 70 mm of usable jaw width and a clamping force of at least 1 kN with a hand force no greater than 100 N at the handle. The vise shall retain its clamping force after the handle is released.

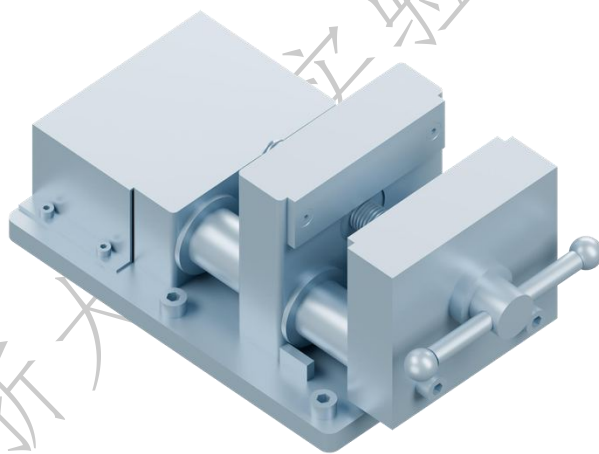


(a) 文本输入片段截图

(b) 功能需求满足率



(c) 参考CAD模型渲染图



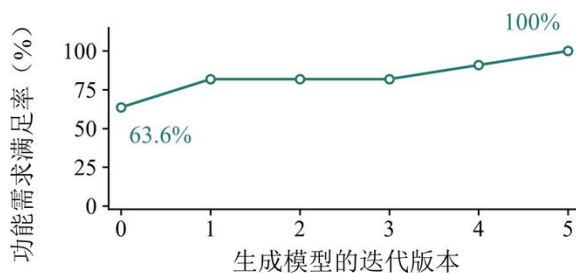
(d) 生成的CAD模型渲染图

图5-9 TRC-02功能需求建模结果

结果分析：TRC-02要求GPT-6设计满足夹持范围、夹紧力和自锁要求的手动台虎钳，并便于安装和维护。GPT-6采用丝杠驱动和双导杆导向，使活动钳口平行移动；迭代中调整止推垫片、限位件及钳口紧固结构，消除了活动部件的局部干涉。最终设计在100 N手力下的计算夹紧力约为1.15 kN，钳口和软垫均可单独更换。功能需求满足率由77.8%提高至100%，见图5-9。台虎钳的36个零件均通过实体有效性检测。

TRC-03（中等）

Design a one-piece cast-iron casing for a single-stage end-suction pump conveying clean water at 10–60 °C. The application has a nominal pump duty of 500 m³/h at 30 m total head. The matching rotor is fixed: its closed impeller has a 650 mm outside diameter, 100 mm overall axial width, 250 mm inlet opening, and 60 mm outlet width, and rotates clockwise at 730 rpm when viewed from the suction side.



(a) 文本输入片段截图

(b) 功能需求满足率



(c) 参考CAD模型渲染图



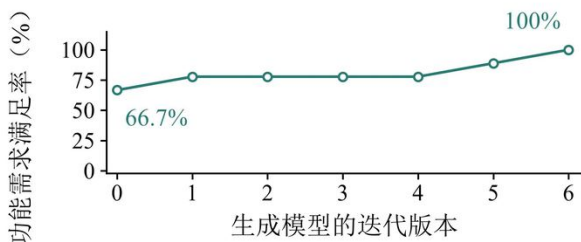
(d) 生成的CAD模型渲染图

图5-10 TRC-03功能需求建模结果

结果分析：TRC-03要求GPT-6围绕指定叶轮设计一体式泵壳，满足导流、承压、管路连接和转子拆卸需求。GPT-6建立了连通的蜗壳流道和后部拆卸开口，并在迭代中调整出口形状、加厚局部承压结构、外移底脚安装孔。最终模型为法兰螺栓和底脚紧固工具留出了操作空间，转子可从后方抽出，水头损失与承压计算均符合给定工况。功能需求满足率由63.6%提高至100%，见图5-10。最终泵壳为单一实体，通过实体有效性检测。

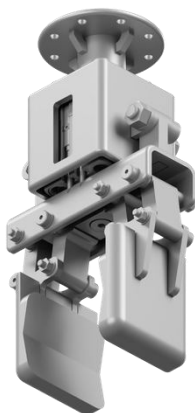
TRC-04（中等）

Design a pneumatic two-finger gripper for transferring dry aluminium workpieces weighing up to 1 kg. Each workpiece has opposing flat gripping areas of at least 30×30 mm, separated by 20–60 mm, with its centre of mass within 20 mm of the midpoint between the contact areas.

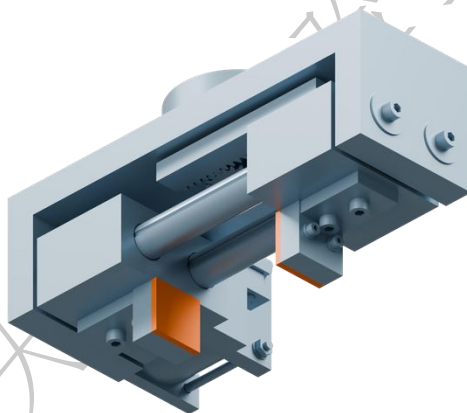


(a) 文本输入片段截图

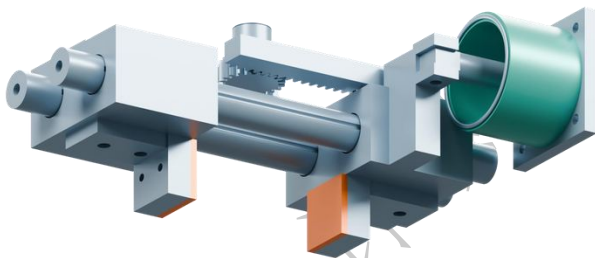
(b) 功能需求满足率



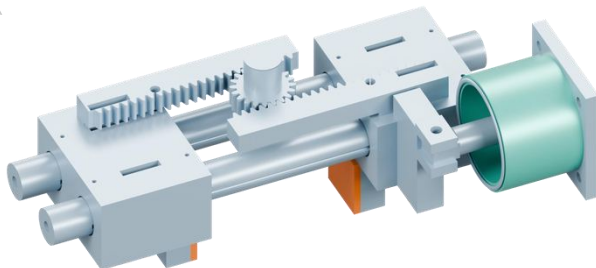
(c) 参考CAD模型渲染图



(d) 生成模型：橙色为夹持软垫



(e) 移去外框后的夹指与导向轴



(f) 内部传动：绿色为气缸筒

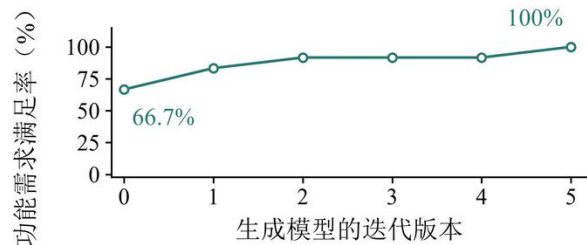
图5-11 TRC-04功能需求建模结果

结果分析：TRC-04要求GPT-6设计能够稳定搬运1 kg工件的气动双指夹爪，并兼顾开口范围与工件表面保护。GPT-6选用了气缸驱动、齿轮齿条同步的平行夹持方案，两根导向轴约束夹指运动，外框支承导向轴和传动部件，因此形成了不同于参考连杆机构的横向框架。两只夹指位于框架下方，橙色软垫直接接触工件。该方案的净开口为71 mm，外形为 $229 \times 117 \times 99$ mm，接近给定尺寸上限。GPT-6通过迭代完善导向和支承结构，将功能需求满足率由66.7%提高至100%，见图5-11。夹爪的89个零件均通过实体有效性检测。

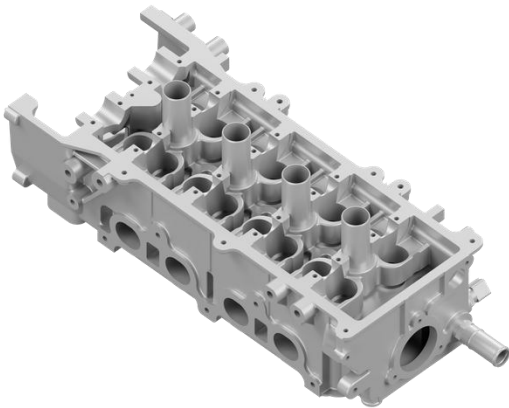
TRC-05（复杂）

Design a one-piece aluminium cylinder-head body for a naturally aspirated, water-cooled, inline-four development engine with a 73 mm bore, 82.8 mm stroke, and 79 mm cylinder spacing. The fixed block interface has a flat deck and ten M10 attachment points arranged in two parallel rows 100 mm apart; each row contains five equally spaced points at 79 mm pitch, centred on the four-cylinder group.

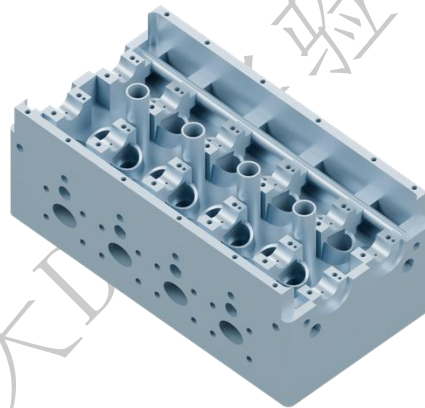
(a) 文本输入片段截图



(b) 功能需求满足率



(c) 参考CAD模型渲染图



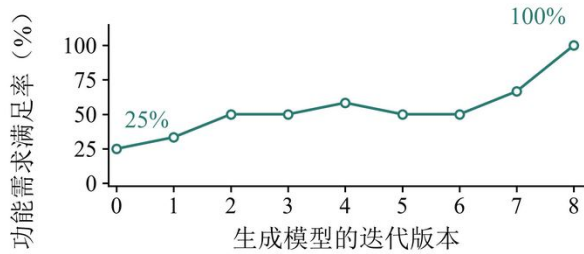
(d) 生成的CAD模型渲染图

图5-12 TRC-05功能需求建模结果

结果分析：TRC-05要求GPT-6设计兼顾配气、承压、冷却、润滑和检修空间的一体式缸盖。GPT-6在迭代中加厚孔道之间的隔壁，完善冷却入口，并打开凸轮轴拆卸通道，功能需求满足率由66.7%提高至100%。最终模型为四个气缸分别设置了气门和火花塞安装孔，进排气道、水套与润滑通道相互隔离，气门和凸轮轴具有所需的运动与拆装空间，见图5-12。最终缸盖通过实体有效性检测。

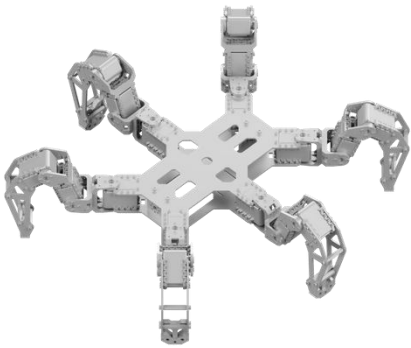
TRC-06（复杂）

Design a six-legged, battery-powered chassis for indoor inspection on dry, hard floors, using 0.5 as the minimum friction coefficient between the feet and ground. Carry a centrally mounted 1 kg inspection module measuring $100 \times 100 \times 60$ mm, with its centre of mass 30 mm above the mounting face and four M4 attachment points on an 80 mm square pattern.

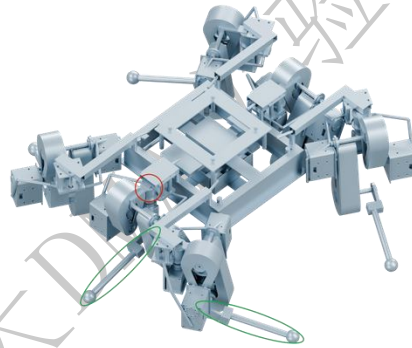


(a) 文本输入片段截图

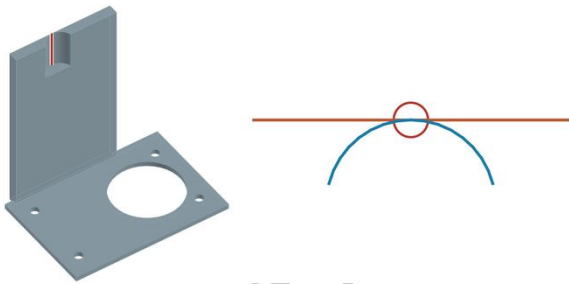
(b) 功能需求满足率



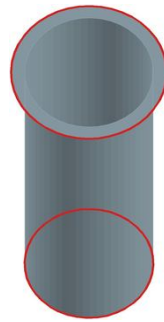
(c) 参考CAD模型渲染图



(d) 生成模型：红圈对应(e)，绿圈为支腿



(e) 电机支承件：实体边界自交



(f) 开放壳：红色边界没有相邻面

图5-13 TRC-06功能需求建模结果

结果分析：TRC-06要求GPT-6设计满足承载、行走、越障和续航需求的六足底盘。GPT-6通过减重和调整腿部运动空间，将功能需求满足率由25%提高至100%。图中绿圈标出向外伸展的小腿与球形足端；六个足端均与水平地面接触，重心投影位于支撑区域内，静力计算表明该布局能够支撑机身及1 kg巡检模块。底盘质量约2.94 kg，计算续航为23.25 min。907个实体中904个通过有效性检测，3个电机支承件因实体边界自交未通过：孔壁与外侧面沿线相切，但相切线并非两面的公共边。图5-13(d)的红圈标出其中一个支承件在整机中的位置，其局部自交缺陷见图5-13(e)。模型另含36个独立开放壳，整机未通过实体有效性检测。

5.3.2 扩展实验结果

扩展实验共包含6项测试。S1和S2分别采用七缸星型发动机与电动直升机机械总成样本，输入为对应参考模型的几何要求或功能需求文本，两项测试均开启网络搜索。S3关闭网络搜索，对TC-05、TC-06、TRC-05和TRC-06进行测试。评测方将装配体拆分为单零件STEP后逐一检测实体有效性，表中分数表示通过检测的实体数与实体总数，检测异常归为未通过。

（1）总体概览

两项压力测试的复杂装配体建模结果见表5-12和表5-13。在S1中，GPT-6完成了七缸径向布置、主副连杆、配气机构及进排气管路的建模，最终模型满足全部12项几何要求。在S2中，GPT-6完成了直升机的主要部件布置及质量、续航设计，但旋翼变距拉杆与倾斜盘断开，云台电机也未固定到支板上。最终方案满足9项功能需求，飞行操纵、横风调姿及相机指向保持3项未满足。两个模型均未通过实体有效性检测：S1的394个实体中385个通过、9个未通过；S2的188个实体中187个通过、1个未通过。

S3无网络搜索实验的结果见表5-14和表5-15。几何要求建模中，TC-05和TC-06的形状、尺寸及空间关系均符合输入要求，两个模型均通过实体有效性检测。功能需求建模中，GPT-6完成了TRC-05缸盖的孔道、配气机构安装空间、承压及冷却结构设计，也完成了TRC-06底盘的载荷安装、行走、越障和续航设计，两例均满足全部12项功能设计要求。但底盘腿部还存在护管延伸至无电气部件的足端、末端悬空的问题。两例均未通过实体有效性检测：TRC-05的局部面边界环存在多点接触；TRC-06的1个足端连接销壳体方向错误，其余348个实体通过检测。各测试样本的用时见表5-16。

表5-12 S1几何要求建模压力测试结果汇总

样本ID	复杂度	实体有效性	几何要求满足情况	测试结果
TC-07	复杂	未通过 385/394	12项全部满足	七缸布置及主要机构符合要求，主连杆存在非流形边，进气管存在曲面自交

表5-13 S2功能需求建模压力测试结果汇总

样本ID	复杂度	实体有效性	功能需求满足情况	测试结果
TRC-07	复杂	未通过 187/188	9项满足，3项未满足	变距拉杆与倾斜盘断开，云台缺少固定连接；下部整流罩未通过实体有效性检测

表5-14 S3几何要求建模无网络实验结果汇总

样本ID	复杂度	实体有效性	几何要求满足情况	测试结果
TC-05	复杂	通过	12项全部满足	缸体为完整单实体，缸孔、斜孔与水套布局符合要求
TC-06	复杂	通过	12项全部满足	气缸、曲柄连杆和配气机构位置正确，整体装配完整

表5-15 S3功能需求建模无网络实验结果汇总

样本ID	复杂度	实体有效性	功能需求满足情况	测试结果
TRC-05	复杂	未通过 0/1	12项全部满足	燃烧室、气道和水套相互隔离，局部面边界环存在多点接触
TRC-06	复杂	未通过 348/349	12项全部满足	腿外侧护管形成多余悬空分支；1个足端连接销壳体方向错误

表5-16 文本到CAD扩展实验各样本测试用时

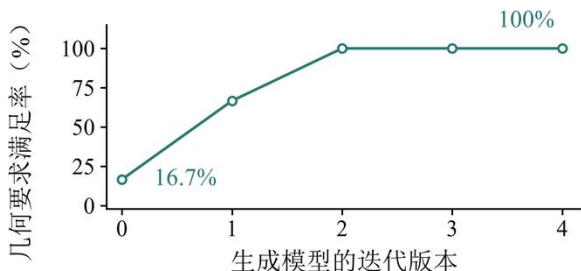
样本ID	S1_TC-07	S2_TRC-07	S3_TC-05	S3_TC-06	S3_TRC-05	S3_TRC-06
用时	40分钟	1时8分钟	48分钟	45分钟	44分钟	1时51分钟

(2) 用例分析

① S1几何要求建模压力测试

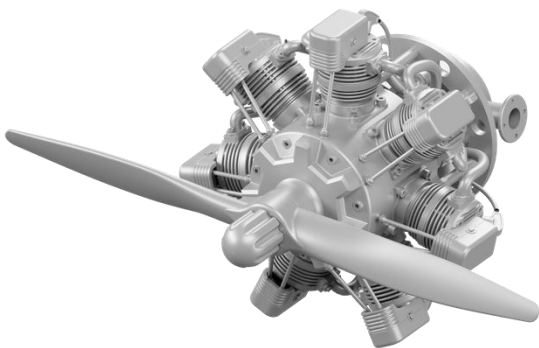
TC-07（复杂）

Create a complete single-row, seven-cylinder radial-engine assembly approximately 280 mm long, with a 440 mm propeller span. The cylinder assemblies, including rocker covers, have an overall envelope diameter of approximately 240 mm. Seven cylinders surround the main shaft at equal angles in a common transverse plane, with one upright. The central crankcase is approximately 100 mm in diameter and 67 mm long.

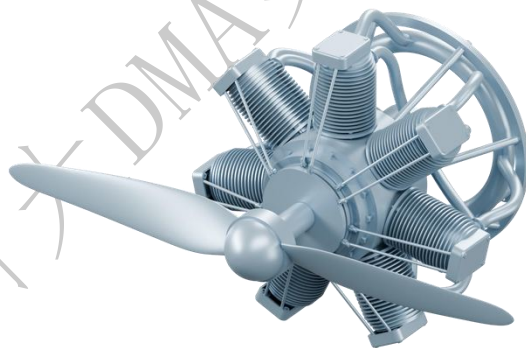


(a) 文本输入片段截图

(b) 几何要求满足率



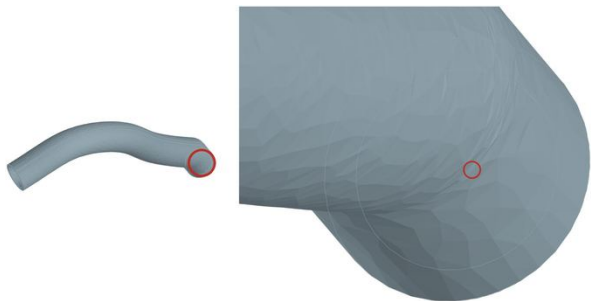
(c) 参考CAD模型渲染图



(d) 生成的CAD模型渲染图



(e) 主连杆：红边被面边界重复使用



(f) 进气管：红圈为曲面自交位置

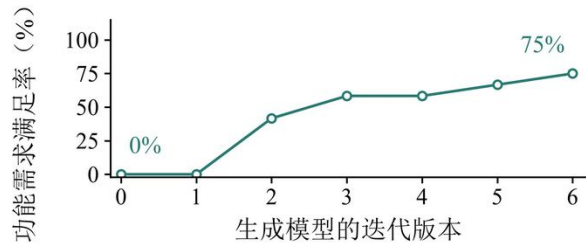
图5-14 TC-07几何要求建模压力测试结果

结果分析：TC-07要求GPT-6完成七缸星型发动机及其主要机构、进排气管路的建模与装配。GPT-6逐步补齐部件，并调整安装环和桨叶截面，几何要求满足率由16.7%提高至100%。最终模型保留了七缸径向布置及主副连杆关系，但主连杆局部存在非流形连接，同一条边在面边界中被使用4次；7根进气管的管壁曲面局部折返并与自身相交，见图5-14。后部进气壳体也未通过参数曲线检测。394个实体中385个通过、9个未通过，整机未通过实体有效性检测。

② S2功能需求建模压力测试

TRC-07（复杂）

Design a compact electric helicopter for visual inspection with a 150 g camera module measuring $60 \times 50 \times 40$ mm. The module has four M3 attachment points on a 40×30 mm rectangular pattern and a centre of mass 20 mm above its mounting face. The complete aircraft, including the camera, battery, and electronics, shall weigh no more than 1.4 kg.

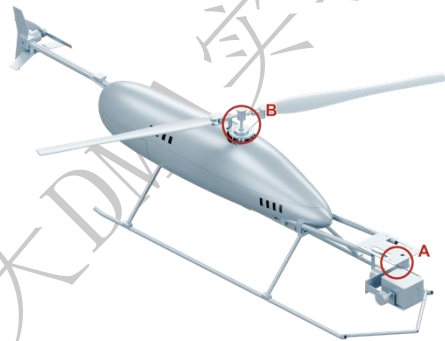


(a) 文本输入片段截图

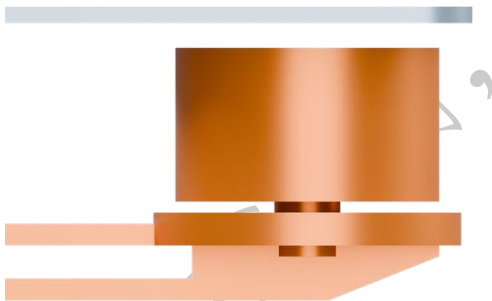
(b) 功能需求满足率



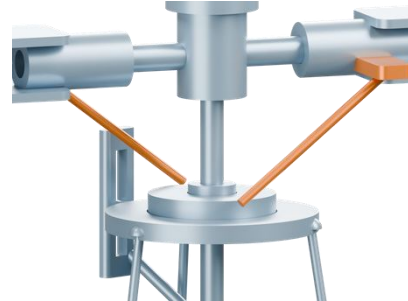
(c) 参考CAD模型渲染图



(d) 生成模型及连接问题位置A、B



(e) A处：云台电机未固定到支板



(f) B处：变距拉杆与倾斜盘断开

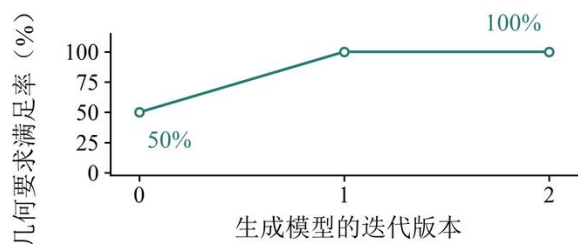
图5-15 TRC-07功能需求建模压力测试结果

结果分析：TRC-07要求GPT-6设计兼顾飞行操纵、悬停续航和相机指向的电动直升机。GPT-6将相机和云台布置在机头外伸支架上，以留出拍摄视野，并通过迭代减重和调整部件位置，将功能需求满足率由0%提高至75%。最终模型的A处云台电机与支板相隔2.25 mm，缺少安装固定；B处变距拉杆与倾斜盘相隔约3.15 mm，无法带动桨叶变距，见图5-15。飞行操纵、横风调姿及相机指向保持3项需求未满足。下部整流罩未通过参数曲线检测，188个实体中187个通过、1个未通过，整机未通过实体有效性检测。

③ S3无网络搜索——几何要求建模

TC-05（复杂）

Create a single solid bare V8 cylinder block approximately 520 mm long, 475 mm wide and 290 mm high. Two cylinder banks form a 90° V, each containing four parallel bores approximately 92 mm in diameter and 110 mm deep. The cylinders repeat at approximately 107.5 mm longitudinal spacing, with the two banks slightly staggered.

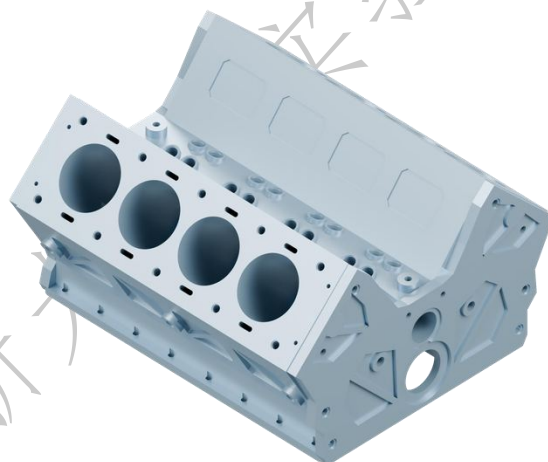


(a) 文本输入片段截图

(b) 几何要求满足率



(c) 参考CAD模型渲染图



(d) 生成的CAD模型渲染图

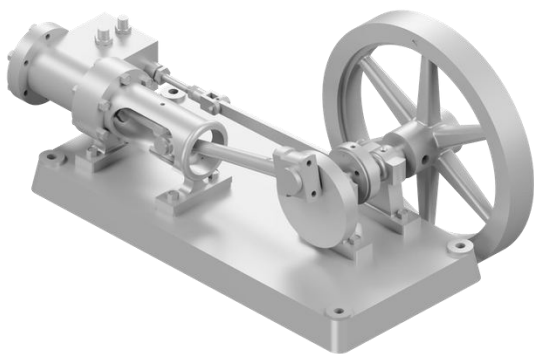
图5-16 TC-05几何要求建模无网络实验结果

结果分析：TC-05要求GPT-6在无网络搜索条件下生成具有完整内部孔道和安装结构的V8缸体。GPT-6先建立两列气缸、水套与纵向通道，再补充挺柱孔、轴承座、底部结合面及外侧加强结构。最终模型的水套保持连通，各腔体之间具有完整隔壁，五道腹板与曲轴轴承座的位置符合要求。几何要求满足率由50%提高至100%，见图5-16。最终缸体为单一实体，通过实体有效性检测。

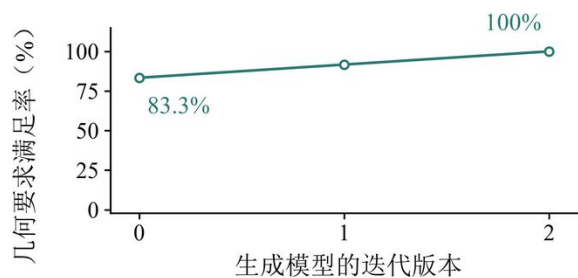
TC-06（复杂）

Create a complete horizontal single-cylinder steam-engine assembly on a rectangular base approximately $170 \times 80 \times 14.5$ mm, with rounded corners, mounting holes and a recessed underside. The cylinder lies along the long direction of the base and extends beyond one end. Its body is 35 mm long, with a 19 mm bore, a narrower barrel approximately 22 mm in diameter and wider end flanges approximately 32 mm in diameter.

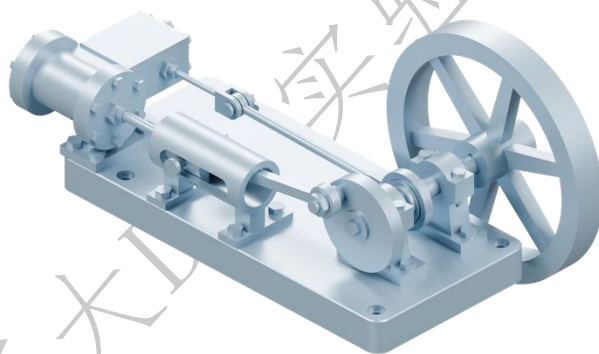
(a) 文本输入片段截图



(c) 参考CAD模型渲染图



(b) 几何要求满足率



(d) 生成的CAD模型渲染图

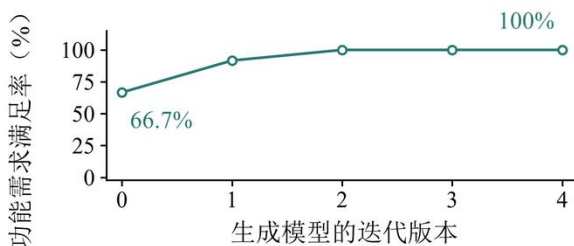
图5-17 TC-06几何要求建模无网络实验结果

结果分析：TC-06要求GPT-6在无网络搜索条件下构建卧式单缸蒸汽机，并完成主传动与配气机构的装配。GPT-6在建立主体机构后补齐紧固件，调整偏心轮、套环和安装螺栓的位置，消除了局部装配干涉。最终模型中，气缸、曲柄连杆和飞轮形成完整的主传动机构，偏心连杆与阀杆的位置符合要求。几何要求满足率由83.3%提高至100%，见图5-17。蒸汽机的110个零件均通过实体有效性检测。

④ S3无网络搜索——功能需求建模

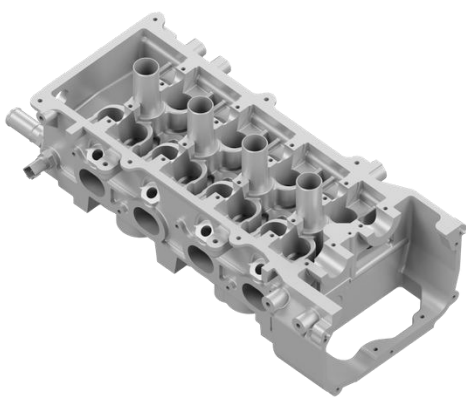
TRC-05（复杂）

Design a one-piece aluminium cylinder-head body for a naturally aspirated, water-cooled, inline-four development engine with a 73 mm bore, 82.8 mm stroke, and 79 mm cylinder spacing. The fixed block interface has a flat deck and ten M10 attachment points arranged in two parallel rows 100 mm apart; each row contains five equally spaced points at 79 mm pitch, centred on the four-cylinder group.

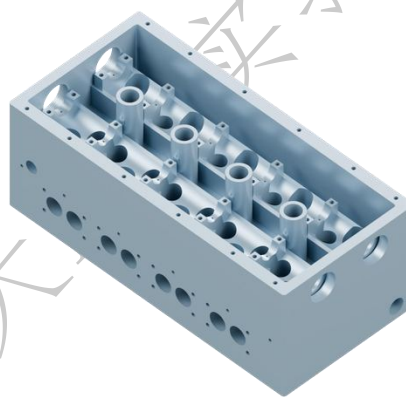


(a) 文本输入片段截图

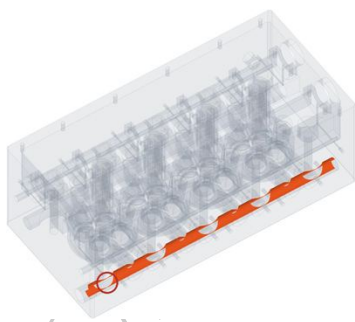
(b) 功能需求满足率



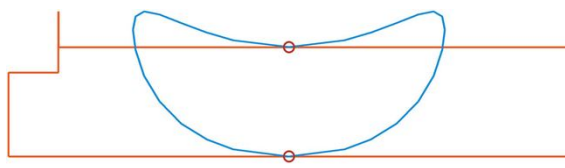
(c) 参考CAD模型渲染图



(d) 生成的CAD模型渲染图



(e) 缸盖内部孔道的交界处



(f) 同一对边界环在两个红圈处相接

图5-18 TRC-05功能需求建模无网络实验结果

结果分析：TRC-05要求GPT-6在无网络搜索条件下完成缸盖的配气、承压、冷却和润滑结构设计。GPT-6通过调整水套与气道之间的隔壁，修正早期设计中两类通道相切的问题，并保留气门运动和配气部件拆卸空间，功能需求满足率由66.7%提高至100%。但最终模型的内部孔道交界处仍存在不合法的面边界：2个面上共有10对边界环分别在两个不同顶点处相接，局部图用不同颜色显示其中一对边界环，并圈出两个接触点，见图5-18。最终缸盖未通过实体有效性检测。

TRC-06（复杂）

Design a six-legged, battery-powered chassis for indoor inspection on dry, hard floors, using 0.5 as the minimum friction coefficient between the feet and ground. Carry a centrally mounted 1 kg inspection module measuring $100 \times 100 \times 60$ mm, with its centre of mass 30 mm above the mounting face and four M4 attachment points on an 80 mm square pattern.



(a) 文本输入片段截图

(b) 功能需求满足率



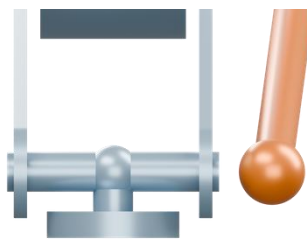
(c) 参考CAD模型渲染图



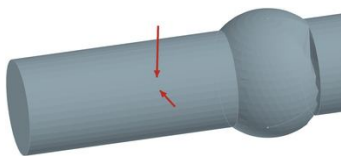
(d) 生成模型：橙色为腿外侧护管



(e) 对应腿部：护管一直延伸至足端



(f) 护管末端与足端连接销分离



(g) 足端连接销的壳体方向错误：红箭头指向实体内部

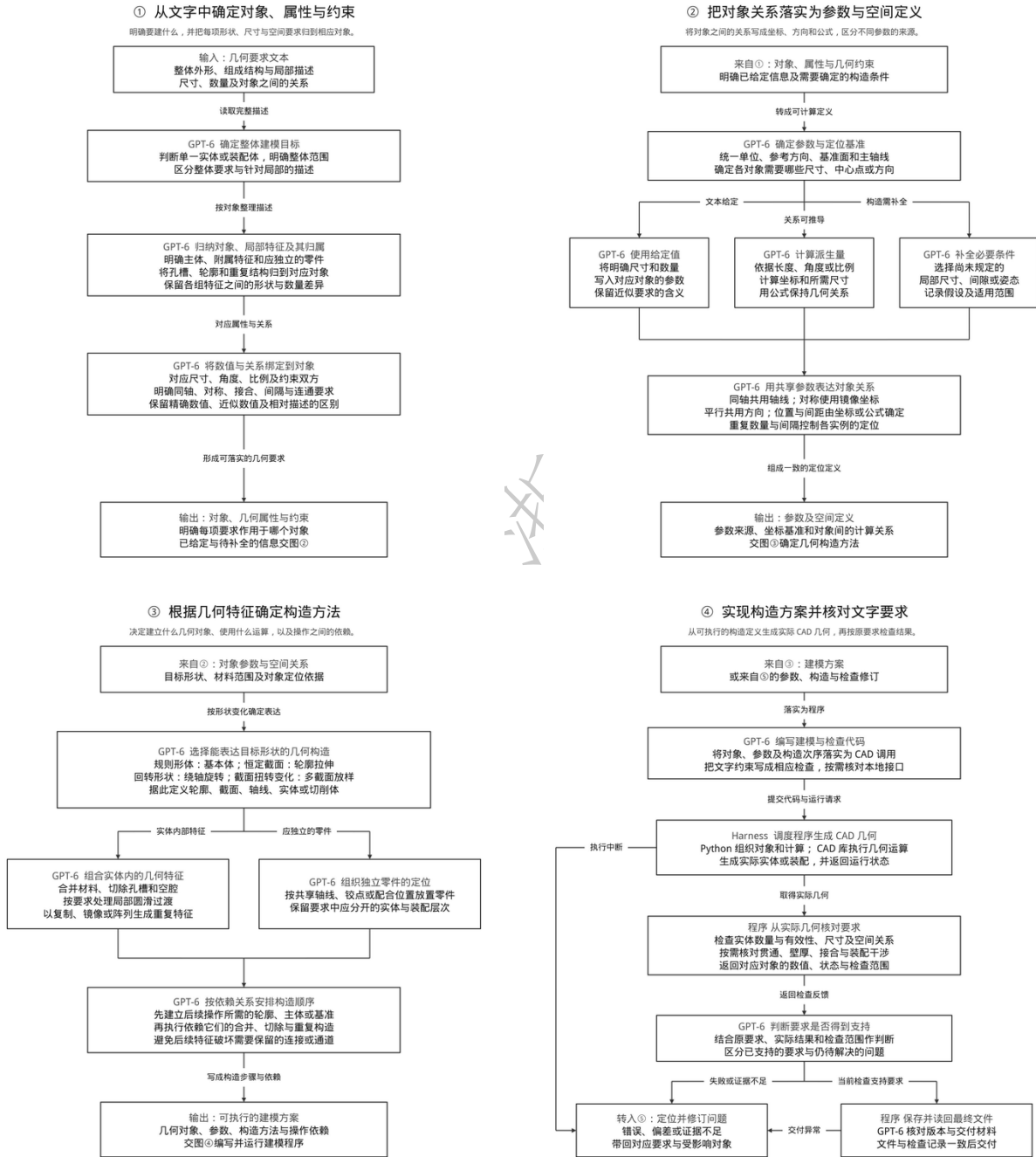
图5-19 TRC-06功能需求建模无网络实验结果

结果分析：TRC-06要求GPT-6设计满足承载、行走和越障需求的六足底盘。GPT-6迭代完善腿部运动，功能需求满足率由33.3%提高至100%。但图中橙色护管一直延伸至没有电气部件的足端，末端悬空，与连接销相隔3 mm，在腿外侧形成多余分支，见图5-19。实体有效性检测发现，1个足端连接销存在壳体方向错误，其局部外表面法向指向实体内部。349个实体中348个通过、1个未通过，底盘未通过实体有效性检测。

5.4 GPT-6推理过程分析

5.4.1 几何要求文本到CAD模型生成的推理过程

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对几何要求文本到CAD模型生成任务可能采用如下推理流程：



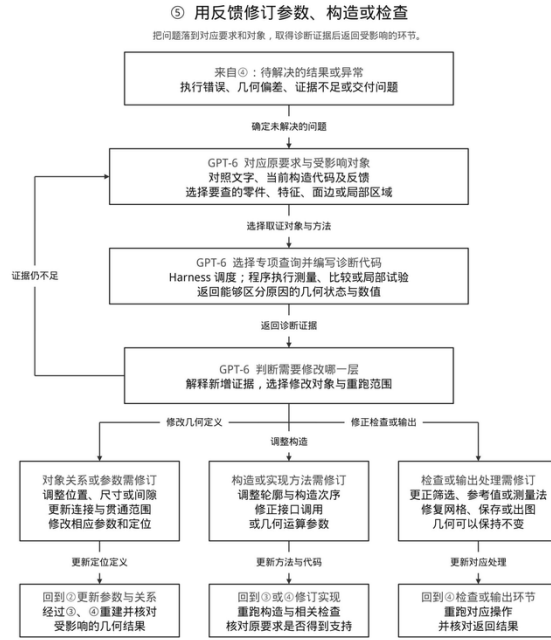
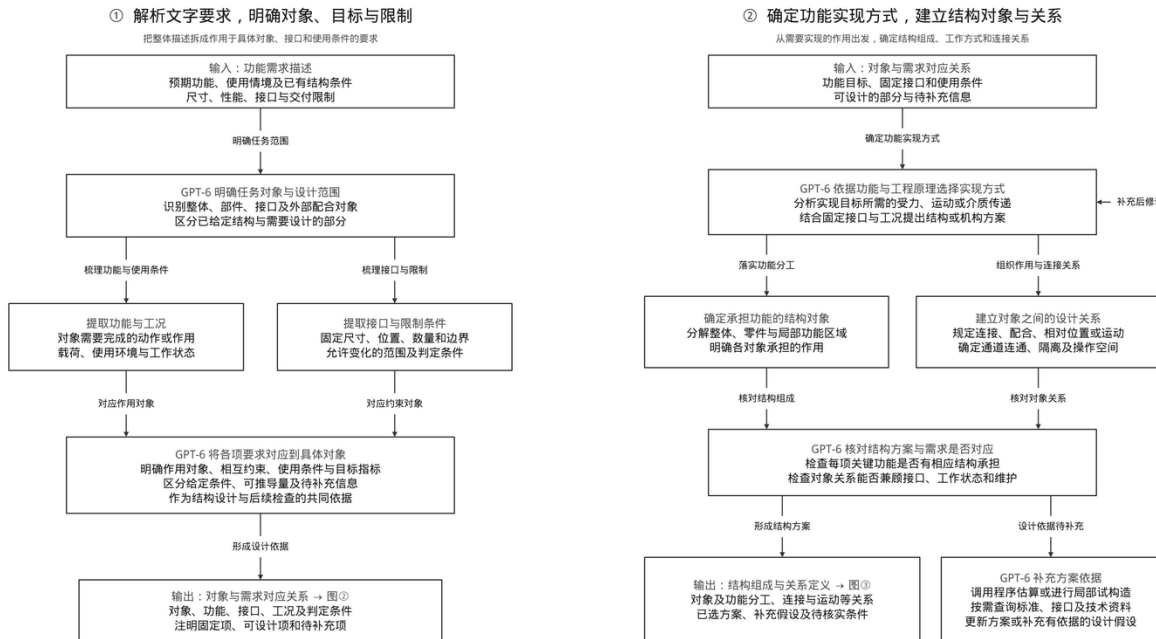


图5-20 几何要求建模推理过程

5.4.2 功能需求文本到CAD模型生成的推理过程

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对功能需求文本到CAD模型生成任务可能采用如下推理流程：



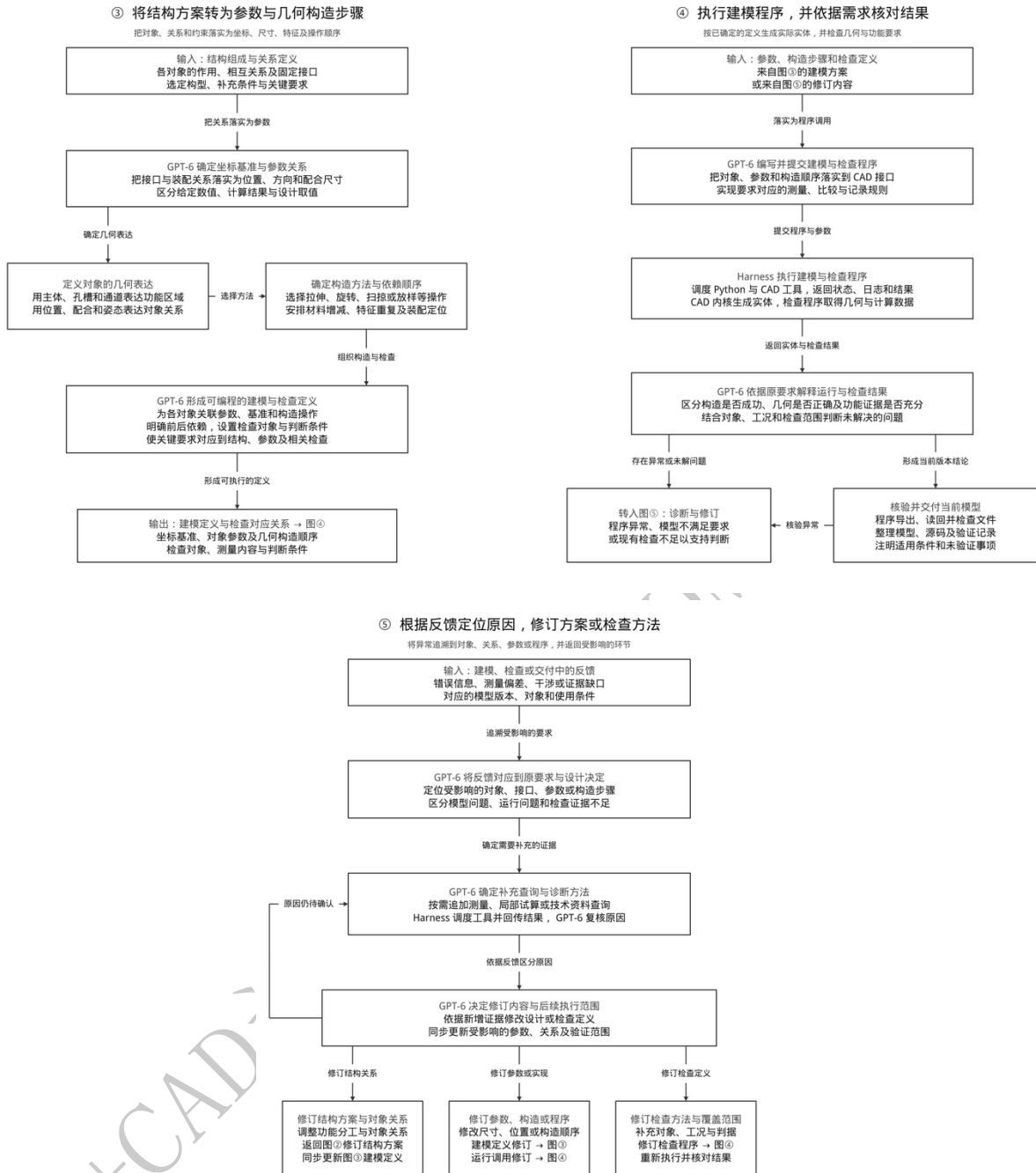


图5-21 功能需求建模推理过程

5.5 能力边界

在本章的文本到CAD模型生成测试中，GPT-6已能够将文本中的几何描述和使用要求转化为结构方案，自主确定尺寸参数与部件关系、规划建模过程，并调用CAD工具生成单零件和装配体模型。GPT-6能够构建重复特征、复杂曲面和内部孔道，并完成多组机构的

布置与装配。建模过程中，GPT-6还能结合受力与运动计算、几何检查等结果，迭代调整结构与参数，补齐遗漏特征并修正部分干涉，逐步提高模型对输入要求的满足程度。

然而，GPT-6在局部细节处理和隐含工程需求的落实方面仍有不足。部分模型虽然满足了文本中明确列出的形状、尺寸或布局要求，但局部过渡、零件固定和传动连接仍不完整，或包含缺乏实际用途的多余结构，影响部件配合与功能实现。对于复杂装配模型，GPT-6生成的部分零件还存在自交、非流形连接或壳体未封闭等几何与拓扑缺陷，导致模型未通过实体有效性检测。GPT-6对这些问题的自主检查与修正仍不够可靠，难以同时保证结构细节合理、机构正常工作和实体有效，生成结果与工程可用水平仍存在差距。

AI+CAD评测，

浙大DMA实验组

6 图像到CAD模型重建/生成

6.1 任务目标与评价标准

图像到CAD模型重建/生成任务主要测试GPT-6对多视角渲染图的理解与三维CAD模型重建能力。主实验采用统一的视图组合和简单中文Prompt，用于获得可比较的重建结果。

本次测试选取6个样本，分别在六视图加轴测图、六视图和三视图三种输入条件下进行建模，共开展18次测试。其中，六视图加轴测图作为主实验，其余两组用于对比视图数量与组合变化对重建结果的影响，各组保持目标模型、Prompt及网络搜索条件一致。在此基础上，选取2个复杂样本开展关闭网络搜索测试，并另选1个复杂壳体开展压力测试，均采用六视图加轴测图输入。主实验与扩展实验合计21次测试，结果分别统计和分析。

表6-1 图像到CAD任务说明

测试要素	说明
测试问题	GPT-6能否根据多视角渲染图，恢复与输入视图相符的CAD模型？视图数量减少后，重建质量如何变化？
核心输入	多视角渲染图
核心输出	可正常读取的B-rep模型
评价标准	检查模型实体有效性，并结合视图轮廓IoU、三维表面HD距离和模型对照，评价几何一致性。

几何一致性分别从视图轮廓和三维表面进行评价，参考STEP仅用于生成后的结果核对。IoU用于衡量生成模型投影与输入渲染图的轮廓重合程度；Hausdorff距离（简称HD）用于衡量生成STEP与参考STEP之间的最大表面偏离，单位为mm，数值越小，记录到的最大偏差越小。计算时，先求参考模型每个采样点到生成模型点云的最近距离，再反向计算生成模型到参考模型的最近距离，取两个方向中最大的距离作为HD。

本次计算先声明单位统一为毫米，按表面积加权采样，采用多个初始姿态和双向ICP进行刚体对齐。对齐过程只允许旋转和平移，不进行缩放或镜像。每个模型使用12,000个点完成配准，再用与配准点分离的180,000个点计算HD，保留全部评估点，本文使用的是有限采样点云的双向最大HD。

6.2 测试用例设计

6.2.1 测试样本选取

图像到CAD模型重建/生成测试样本选取公开的机械零件与装配体。优先考虑结构具有代表性且能够获得对应STEP参考文件的案例。为覆盖不同难度的建模任务，样本在结构复杂程度、局部特征类型及零件数量等方面保持一定差异，涵盖薄板、外壳、带孔支撑件、曲面支架和底盘装配体，并同时包含单零件和多零件装配体建模场景。

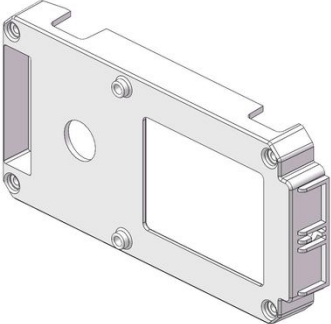
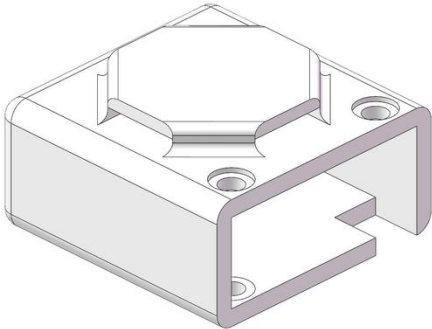
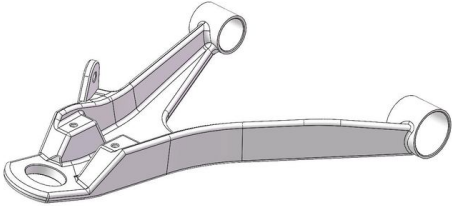
具体而言，在输入视图数量一致的条件下，样本复杂度主要根据零件的几何结构、局部特征之间的空间关系及遮挡程度进行划分。其中多零件样本还进一步考虑零件数量及装配关系。具体划分标准如表6-2所示。

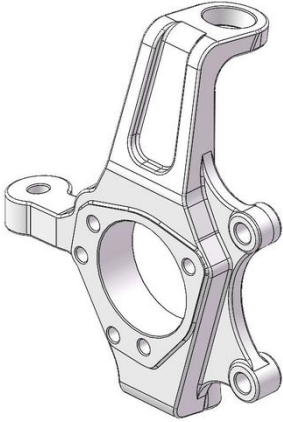
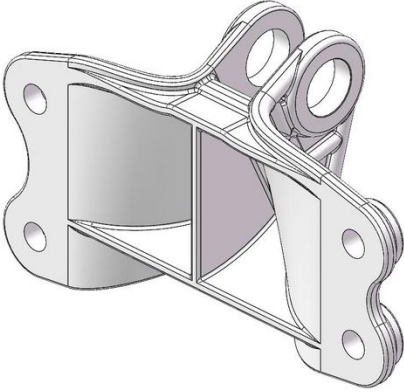
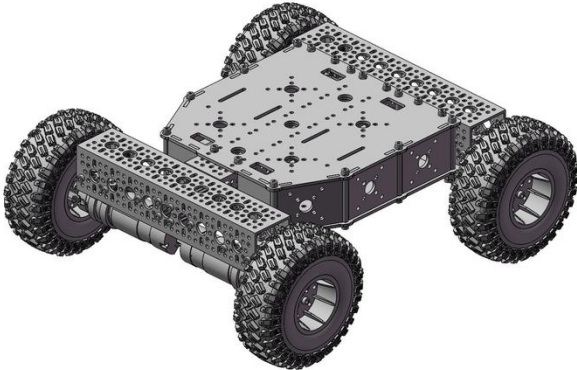
表6-2 图像到CAD模型重建/生成任务样本复杂度分级表

样本复杂度	选取标准	典型表现
简单	渲染图中零件主体较规则，局部特征数量较少，主要轮廓、孔槽及凸台在不同视图中的对应关系较明确，复杂曲面和遮挡结构较少。	主要根据清晰的轮廓和少量局部特征即可判断零件的整体结构，跨视图关联及局部形状推断需求较低。
中等	渲染图中局部细节和空间关系增多，零件包含不同方向的孔、倾斜连接或变截面结构，需要结合多个视图判断特征的形状、朝向及连接关系，但主体仍具有较强规律性。	需要综合多个视图的信息确定局部形状及空间位置，但仍可利用明显的形状规律、重复特征或较明确的连接关系辅助建模。
复杂	渲染图中结构信息较密集，包含复杂曲面、薄壁及相互交汇的局部特征，部分区域存在遮挡，仅凭整体轮廓或重复关系难以确定局部形状；多零件样本还包含较多构件及需要结合其他零件判断的连接结构。	需要综合分散在不同视图中的几何信息，处理曲面变化、结构层次及连接过渡等复杂关系；在多零件场景下，还需结合零件间的位置和对应关系理解局部结构，遮挡进一步增加了判断难度。

依据上述标准，主测试共选取6个样本，包括5个单零件样本和1个多零件装配体样本，简单、中等和复杂样本各2个。具体样本清单如表6-3所示。

表6-3 图像到CAD模型重建/生成任务测试样本清单

样本ID	CAD模型渲染图	样本类型	复杂度	复杂度归类说明
I2C-01		单零件	简单	主体为带翻边的薄板，窗口、圆孔及凸合的布局较规则，主要边界可由正投影视图对应。
I2C-02		单零件	简单	主体由规则壳体和顶部平台构成，开口及导向结构较少，主要轮廓和特征位置在不同视图中较易对应，复杂曲面及跨视图关联需求较少。
I2C-03		单零件	中等	横向衬套、倾斜眼板与加强壁相连，需要结合多个视图判断孔轴方向、眼板倾斜及局部连接关系；但叉形主体和主要支承布局较明确，各部分仍可分解为较清楚的建模特征。

样本ID	CAD模型渲染图	样本类型	复杂度	复杂度归类说明
I2C-04		单零件	中等	中央孔系、安装耳与变截面弯臂相连，需要综合多个视图确定弯臂的空间弯折、截面变化及端部孔位置；但主要结构数量有限，各部分的连接关系较明确。
I2C-05		单零件	复杂	非对称安装耳通过弯曲薄翼与双耳孔连接，翼面前后分布多处筋条，需要结合多个视图判断翼面曲率、局部厚度及筋条交汇处的过渡。各特征相互影响，局部形状难以由简单轮廓或重复规律直接确定。
I2C-06		装配体	复杂	由板件、横梁、车轮和电机支承等多类构件组成，构件数量较多，孔阵与连接结构密集，部分区域存在相互遮挡。除各构件自身的几何建模外，还需综合多个视图判断构件之间的位置及连接关系。

6.2.2 测试条件设计

(1) 主实验设计

主实验使用六个样本的六视图加轴测图组合和统一建模Prompt，不针对样本复杂度及具体结构提供额外提示，要求GPT-6根据输入渲染图完成B-rep模型构建并导出STEP文件。

主实验中，所有任务采用统一Prompt，不针对样本信息及具体结构提供额外提示，要求GPT-6根据输入渲染图完成三维CAD模型构建并输出相应模型文件及预览结果。对于多零件装配体样本，在各零件建模完成后，进一步按照统一要求进行装配验证。主实验保持网络搜索权限开启，以考察GPT-6在正常使用条件下的渲染图到CAD模型重建/生成能力。具体Prompt内容为：“请根据这些渲染图建立对应的CAD模型，并导出STEP文件。”

(2) 扩展实验设计

在主实验基础上，选取具有代表性的样本开展扩展实验，进一步考察渲染图输入信息和网络搜索条件变化对重建/生成结果的影响。扩展实验主要设置两种条件变化：输入视图数量减少、关闭网络搜索权限。各扩展实验均沿用主实验中的对应Prompt，不针对具体样本增加额外提示。扩展实验结果分别与对应样本的主实验结果进行对照。具体设置如表6-4所示。

表6-4 图像到CAD模型重建/生成任务扩展实验设置

设置	输入视图	网络搜索	测试目的
S1视图减少	三视图	开启	考察视图明显减少后GPT-6建模能力。
S2移除轴测图	六视图	开启	考察移除轴测图后GPT-6建模能力。
S3关闭搜索	六视图加轴测图	关闭	考察无法获取外部网络信息时模型的建模表现。
S4压力测试	六视图加轴测图	开启	考察参考模型复杂度过高时GPT-6模型的建模能力。

6.3 测试结果展示

本节先汇总多视图的主实验结果和扩展实验结果，再逐例展示输入渲染图、参考与生成模型、内部迭代曲线及结果分析。

6.3.1 主实验结果

(1) 总体概览

主实验共包含5个单零件样本和1个多零件样本，均导出了STEP文件，并通过拓扑有效性、壳闭合和正体积检查。六视图加轴测图的平均IoU为0.8790—0.9915，HD为9.959—

80.297 mm，总体结果见表6-5。生成模型大多保留了主体外形，但局部结构和实际尺寸仍有偏差。例如I2C-02的IoU达到0.9915，生成模型的外形尺寸却约为参考模型的3.9倍，HD达到64.207 mm。轮廓重合程度与三维尺寸一致性需要分别判断。基本实体检查通过，也不能说明自由曲面、实体连接和装配接口均已正确还原。

表6-5 图像到CAD模型重建/生成主实验结果概览

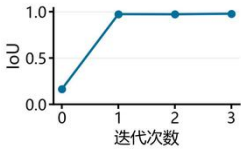
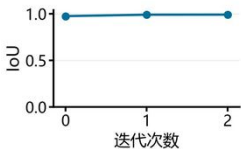
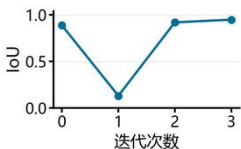
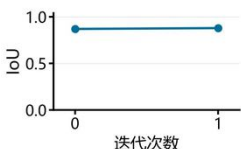
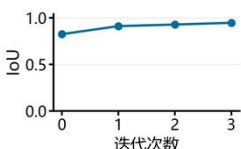
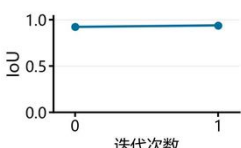
样本ID	难度	实体有效性	六视图加轴测图 平均IoU	HD距离 (mm)	IoU与迭代次数曲线图
I2C-01	简单	通过	0.9788	14.758	
I2C-02	简单	通过	0.9915	64.207	
I2C-03	中等	通过	0.9483	62.401	
I2C-04	中等	通过	0.8790	45.211	
I2C-05	复杂	通过	0.9468	9.959	
I2C-06	复杂	通过	0.9391	80.297	

表6-6 图像到CAD主实验各样本测试用时

样本ID	I2C-01	I2C-02	I2C-03	I2C-04	I2C-05	I2C-06
用时	16分钟	13分钟	26分钟	9分钟	21分钟	3时41分钟

(2) 用例分析

以下按照样本编号逐一展示主实验结果。每个用例包括样例描述、输入渲染图、结果展示和结果分析，重点说明主体结构的还原情况及局部差异。

I2C-01 TrackerOne背板（简单）

输入渲染图

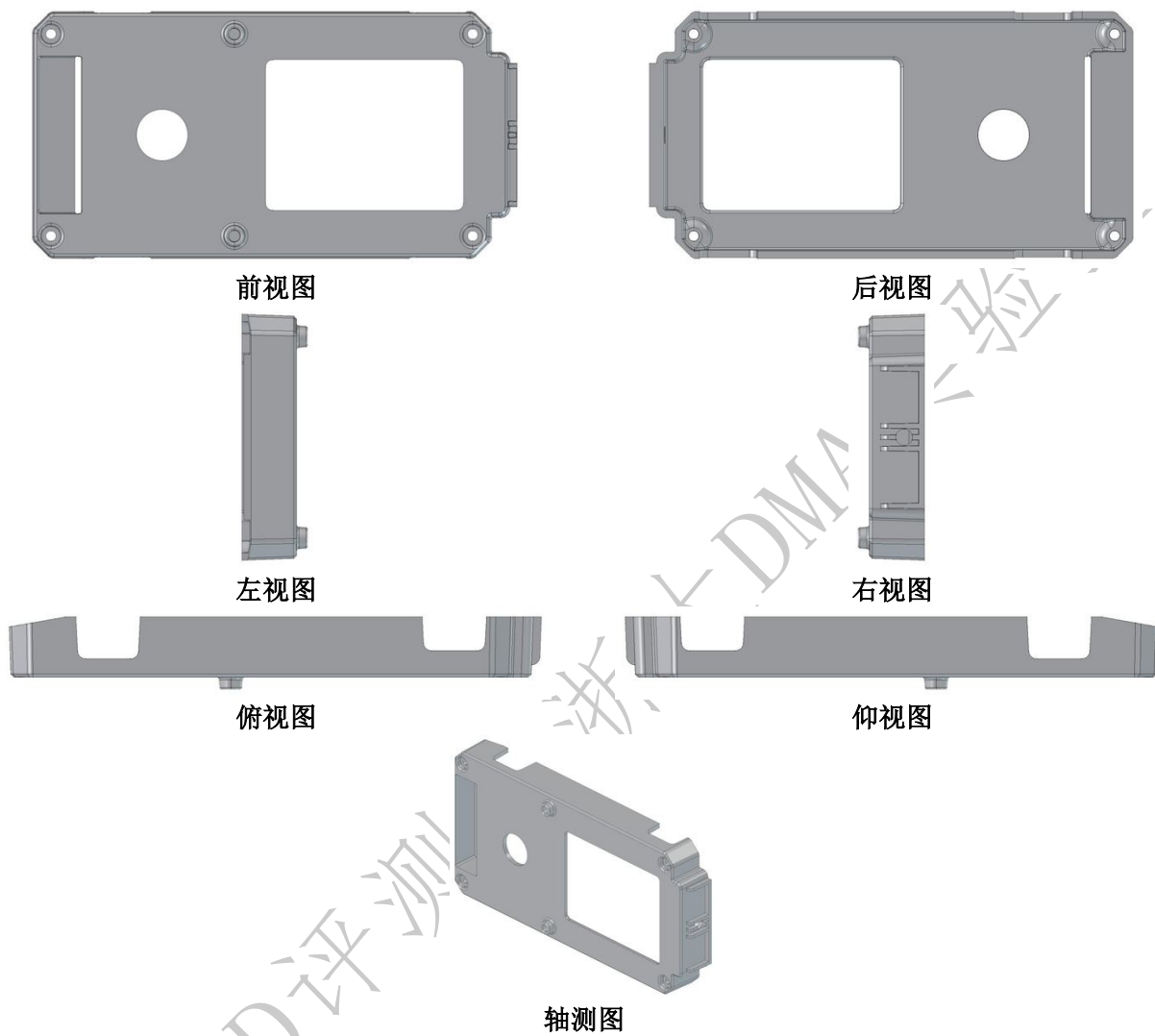
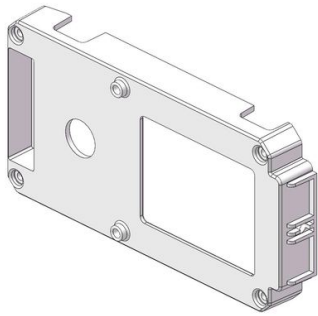
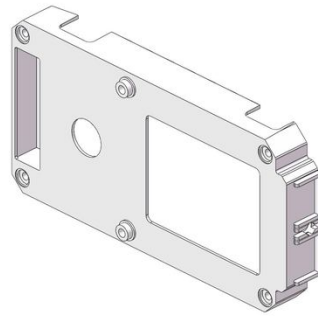


图6-1 I2C-01的七张输入渲染图

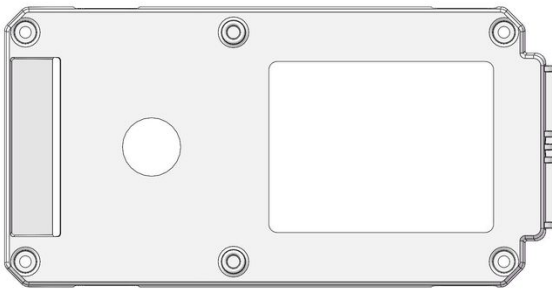
结果展示



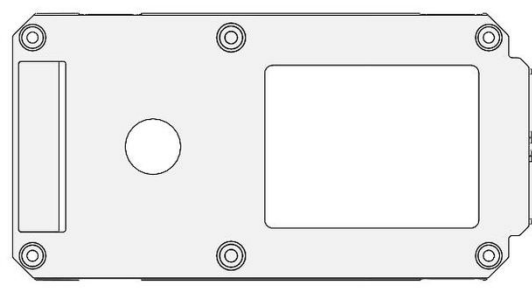
参考CAD模型渲染图（轴测图）



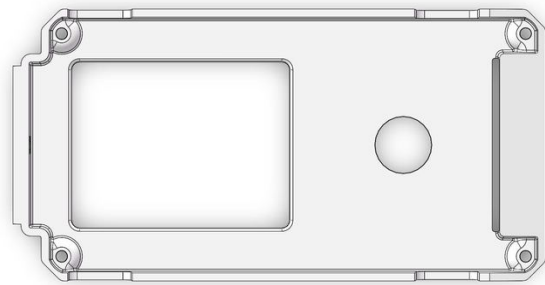
生成的CAD模型渲染图（轴测图）



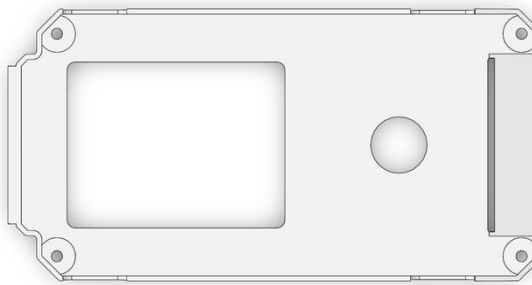
参考CAD模型渲染图（前视图）



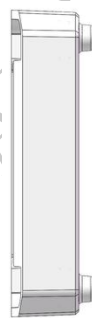
生成的CAD模型渲染图（前视图）



参考CAD模型渲染图（后视图）



生成的CAD模型渲染图（后视图）



参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）

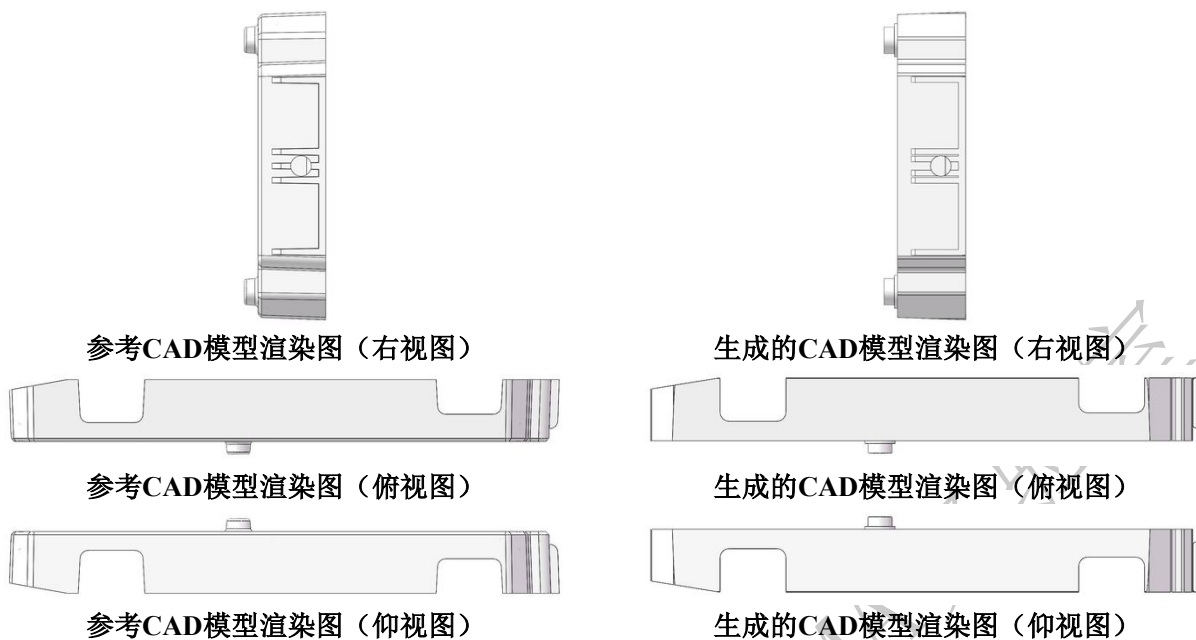


图6-2 I2C-01参考与生成模型的渲染图

结果分析：I2C-01的生成模型通过基本实体检查，窗口、圆孔及周边翻边的总体布局与参考模型接近。图6-2中，侧面卡扣、角部孔座及凸台根部仍存在形状差异，部分小过渡被简化。生成模型沿长边方向约为158.00 mm，参考模型约为132.11 mm，整体尺度偏大；对应HD为14.758 mm。主要结构已还原，但卡扣和孔座轮廓关系到卡接定位与紧固。

各视图的轮廓IoU随迭代次数的变化见图6-3。

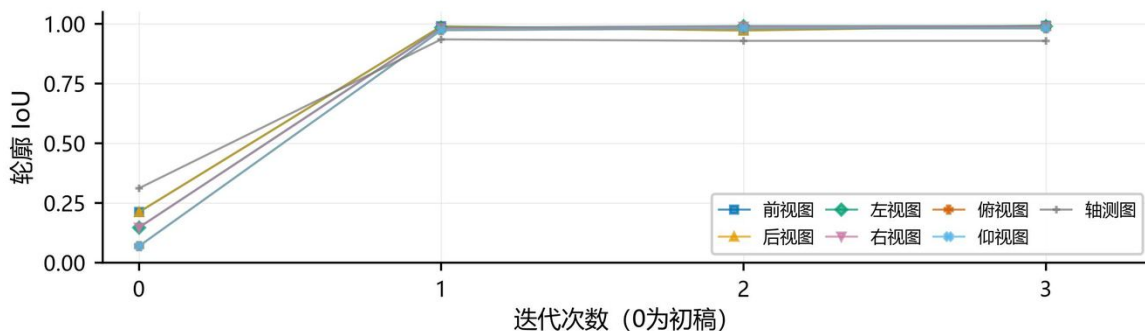


图6-3 I2C-01的迭代次数与IoU曲线

I2C-02 KnowCO2传感器外壳（简单）

输入渲染图

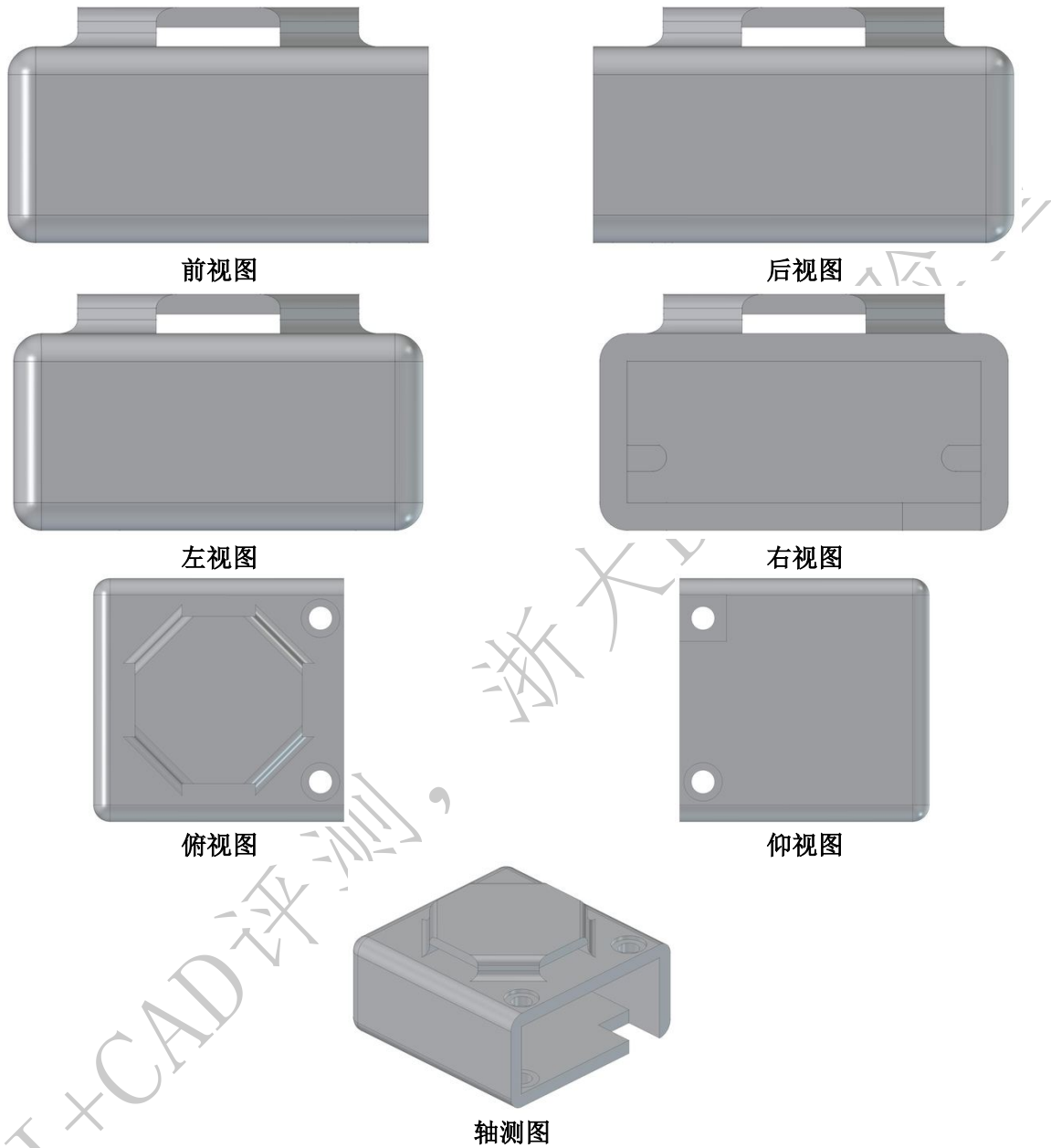
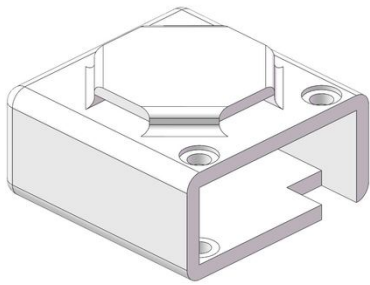
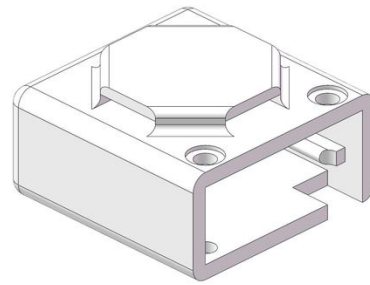


图6-4 I2C-02的七张输入渲染图

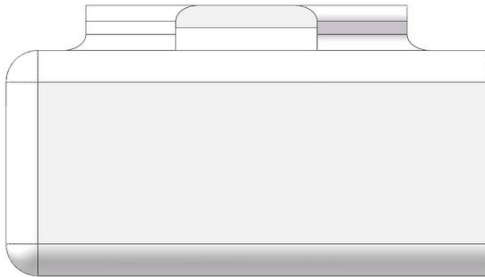
结果展示



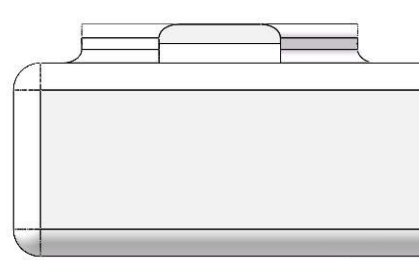
参考CAD模型渲染图（轴测图）



生成的CAD模型渲染图（轴测图）



参考CAD模型渲染图（前视图）



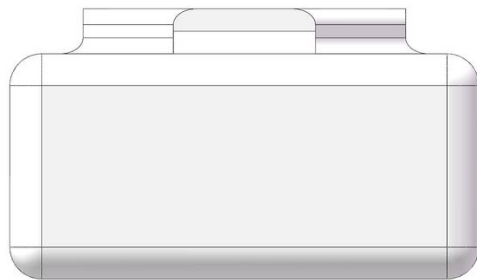
生成的CAD模型渲染图（前视图）



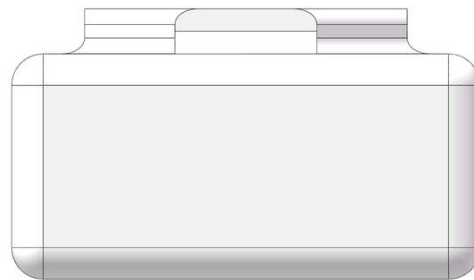
参考CAD模型渲染图（后视图）



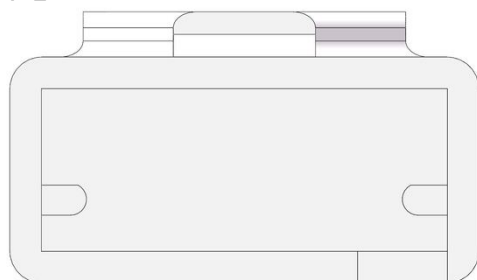
生成的CAD模型渲染图（后视图）



参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）



参考CAD模型渲染图（右视图）



生成的CAD模型渲染图（右视图）

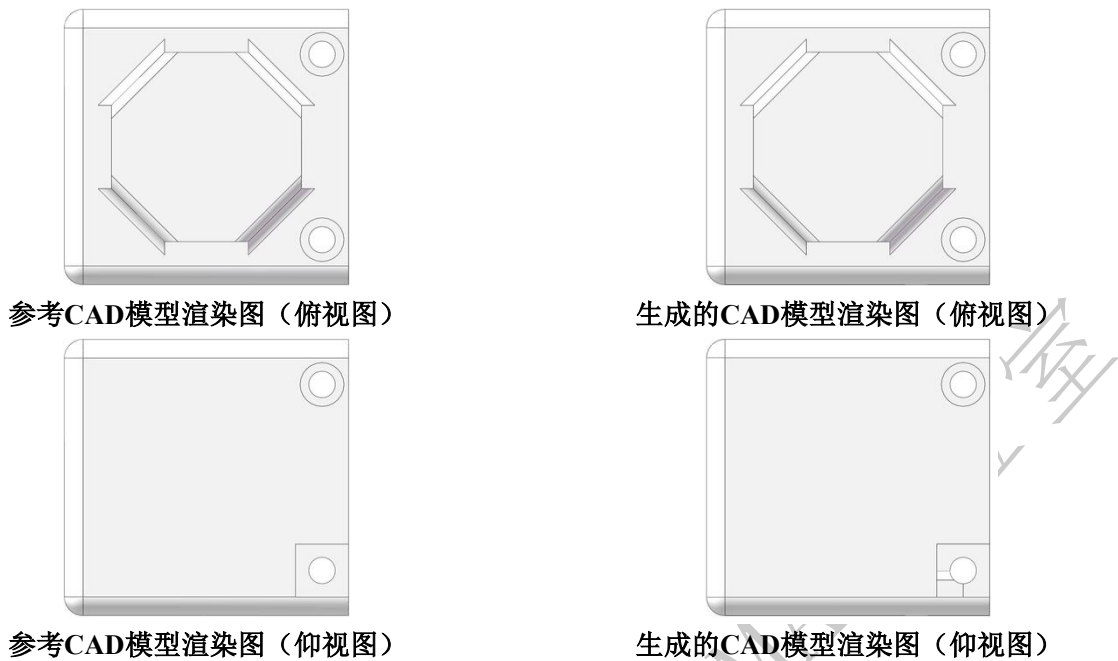


图6-5 I2C-02参考与生成模型的渲染图

结果分析：I2C-02的生成模型通过基本实体检查，壳体开口、顶部平台及两侧孔座的布局与参考模型接近，平台根部和内部导向结构仍有细节差别。其七视图平均IoU为0.9915，但参考模型外形尺寸约为27.18×26.42×15.24 mm，生成模型为107.00×104.00×60.00 mm，整体明显放大，HD为64.207 mm。该样本主要外形比例接近参考模型，绝对尺寸未能还原。

各视图的轮廓IoU随迭代次数的变化见图6-6。

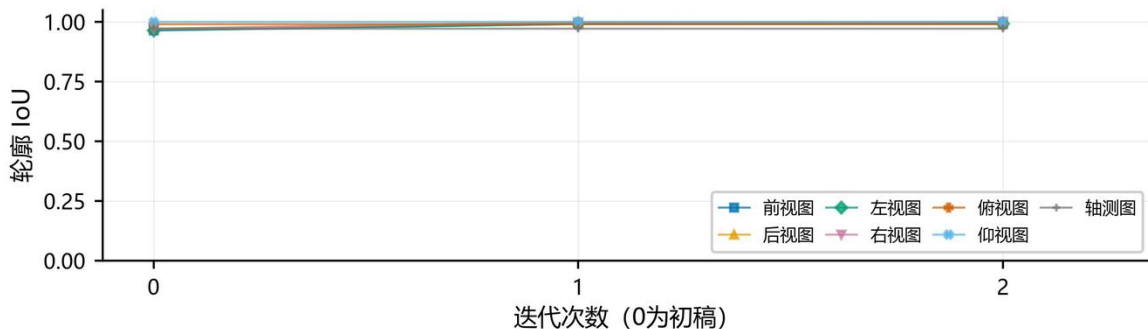


图6-6 I2C-02的GPT内部迭代次数与IoU曲线

I2C-03 C5 Corvette下控制臂（中等）

输入渲染图

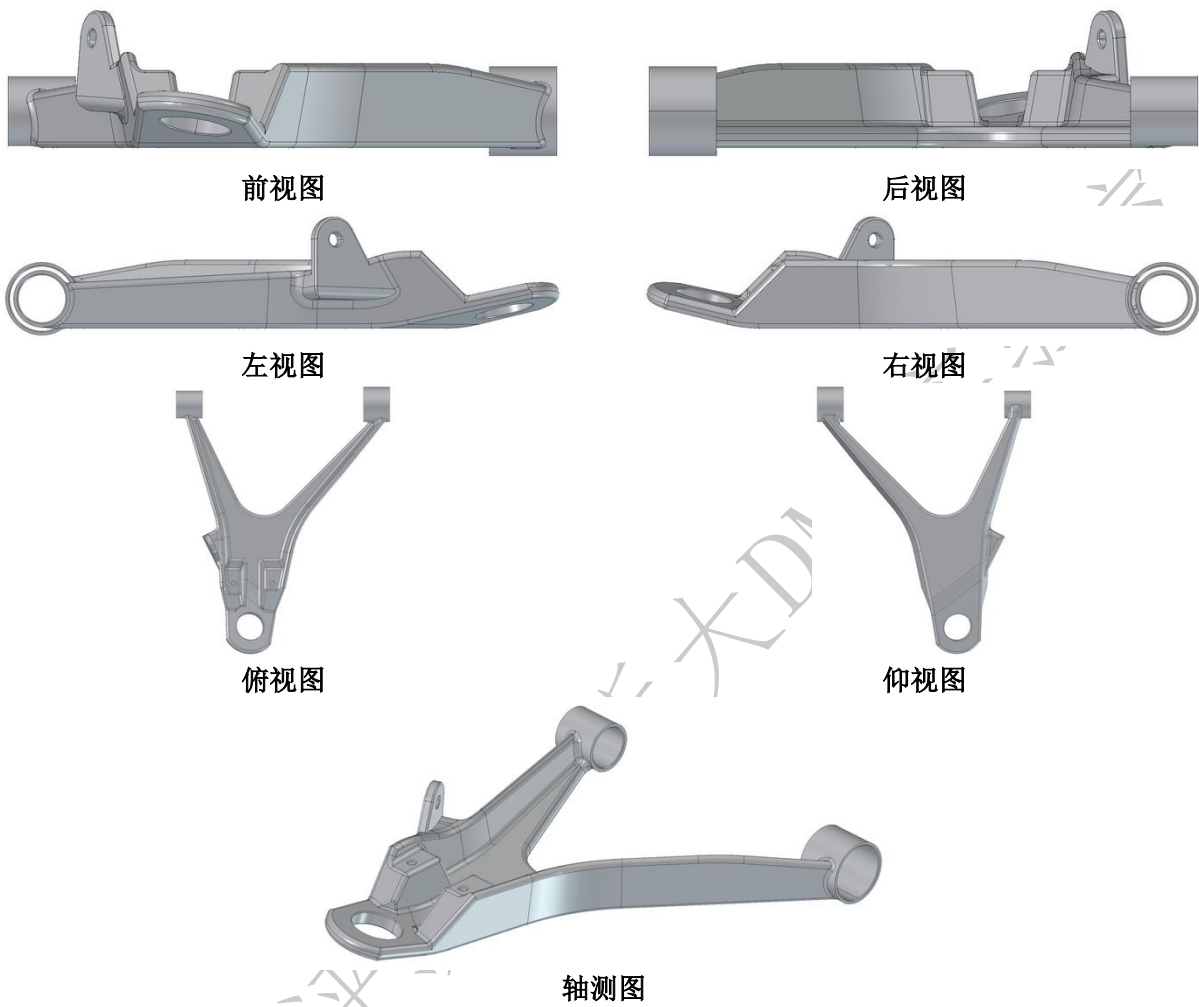
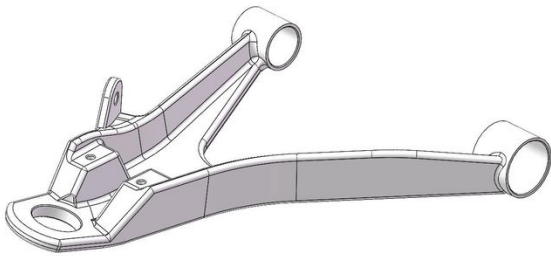
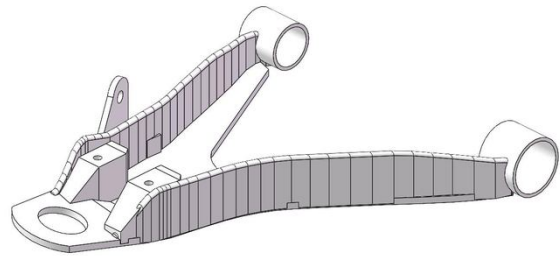


图6-7 I2C-03的七张输入渲染图

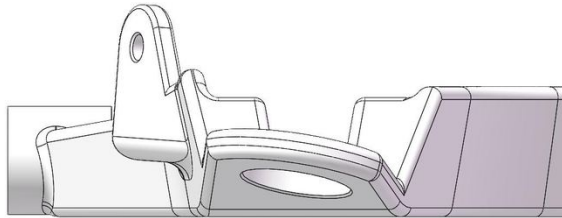
结果展示



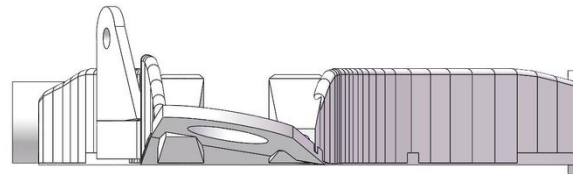
参考CAD模型渲染图（轴测图）



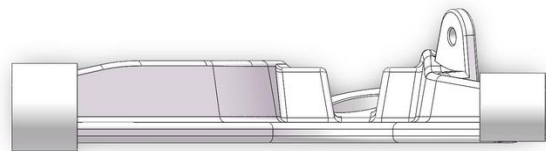
生成的CAD模型渲染图（轴测图）



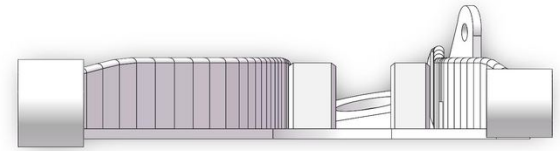
参考CAD模型渲染图（前视图）



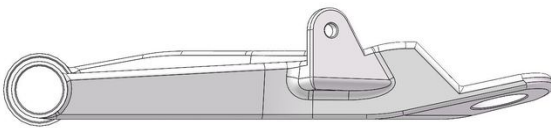
生成的CAD模型渲染图（前视图）



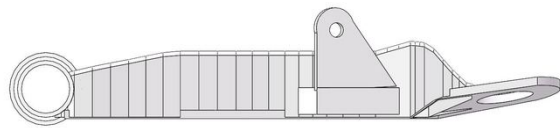
参考CAD模型渲染图（后视图）



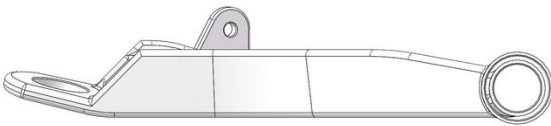
生成的CAD模型渲染图（后视图）



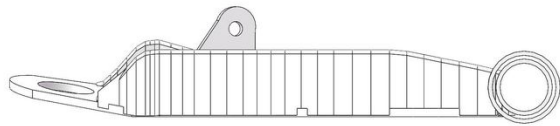
参考CAD模型渲染图（左视图）



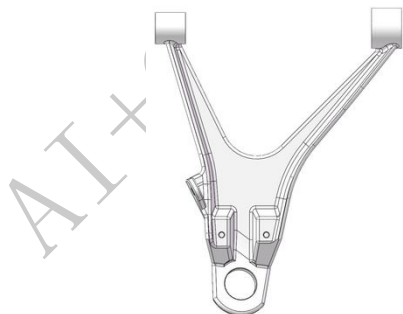
生成的CAD模型渲染图（左视图）



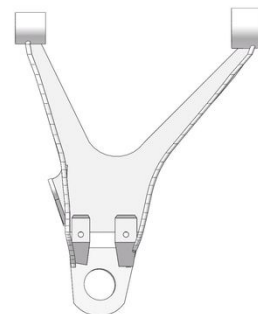
参考CAD模型渲染图（右视图）



生成的CAD模型渲染图（右视图）



参考CAD模型渲染图（俯视图）



生成的CAD模型渲染图（俯视图）



图6-8 I2C-03参考与生成模型的渲染图

结果分析：I2C-03的生成模型实体有效，叉臂、两端衬套及前端耳板的相对位置基本正确。但叉臂侧壁采用多截面直纹放样近似，原有光顺曲面被分段曲面替代，根部圆角缺失，局部还形成了台阶，细节如下图所示。生成模型包围盒最大边长约为394.76 mm，参考模型为488.73 mm，整体尺寸偏小，HD为62.401 mm。该样本的主要问题是曲面过渡和截面连接不准确。

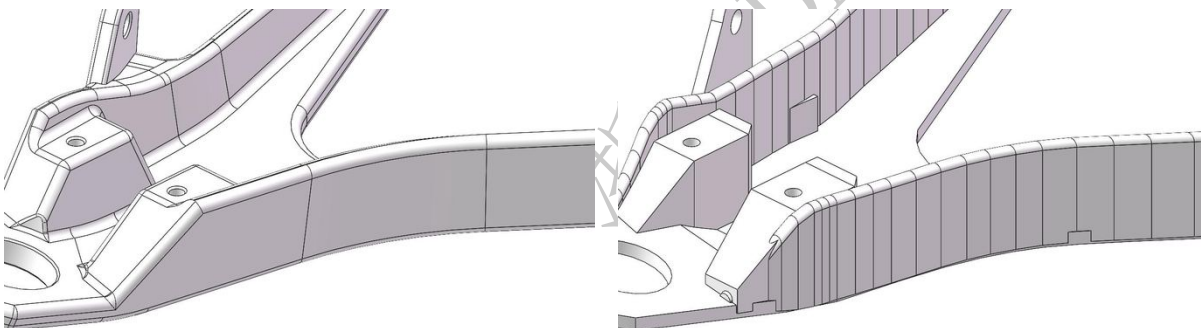


图6-9 I2C-03叉臂侧壁及根部过渡对比

各视图的轮廓IoU随迭代次数的变化见图6-10。

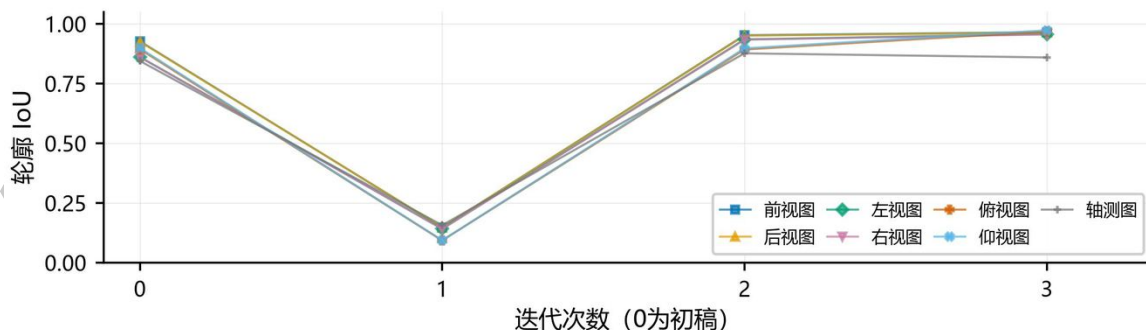


图6-10 I2C-03的GPT内部迭代次数与IoU曲线

I2C-04 C5 Corvette转向节（中等）

输入渲染图

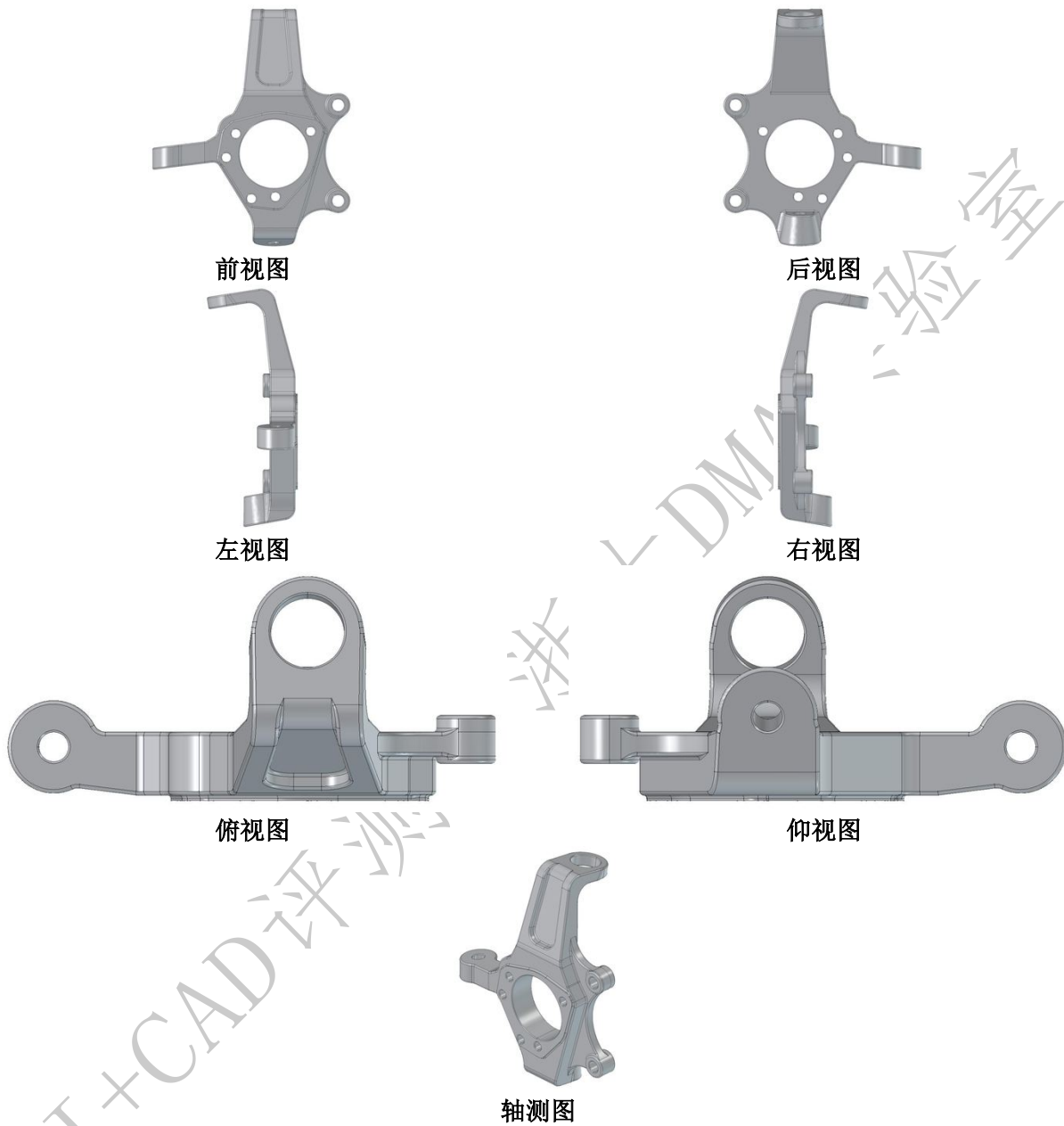
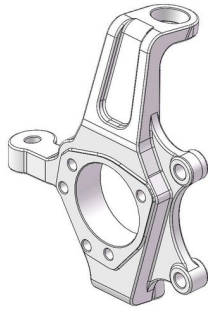
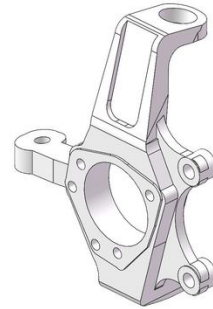


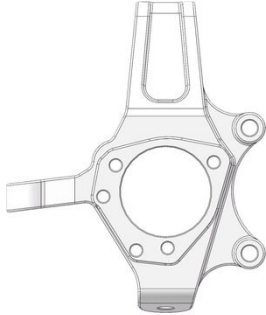
图6-11 I2C-04的七张输入渲染图



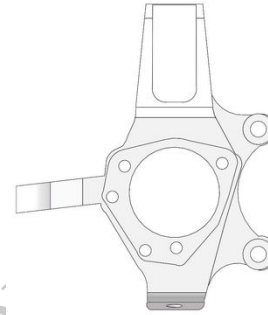
参考CAD模型渲染图（轴测图）



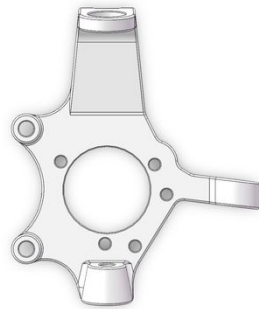
生成的CAD模型渲染图（轴测图）



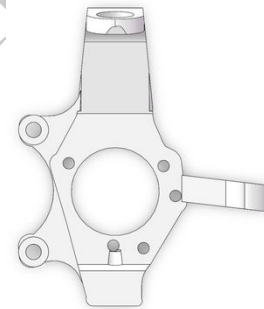
参考CAD模型渲染图（前视图）



生成的CAD模型渲染图（前视图）



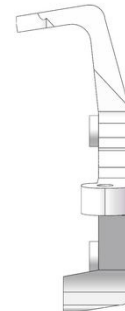
参考CAD模型渲染图（后视图）



生成的CAD模型渲染图（后视图）



参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）

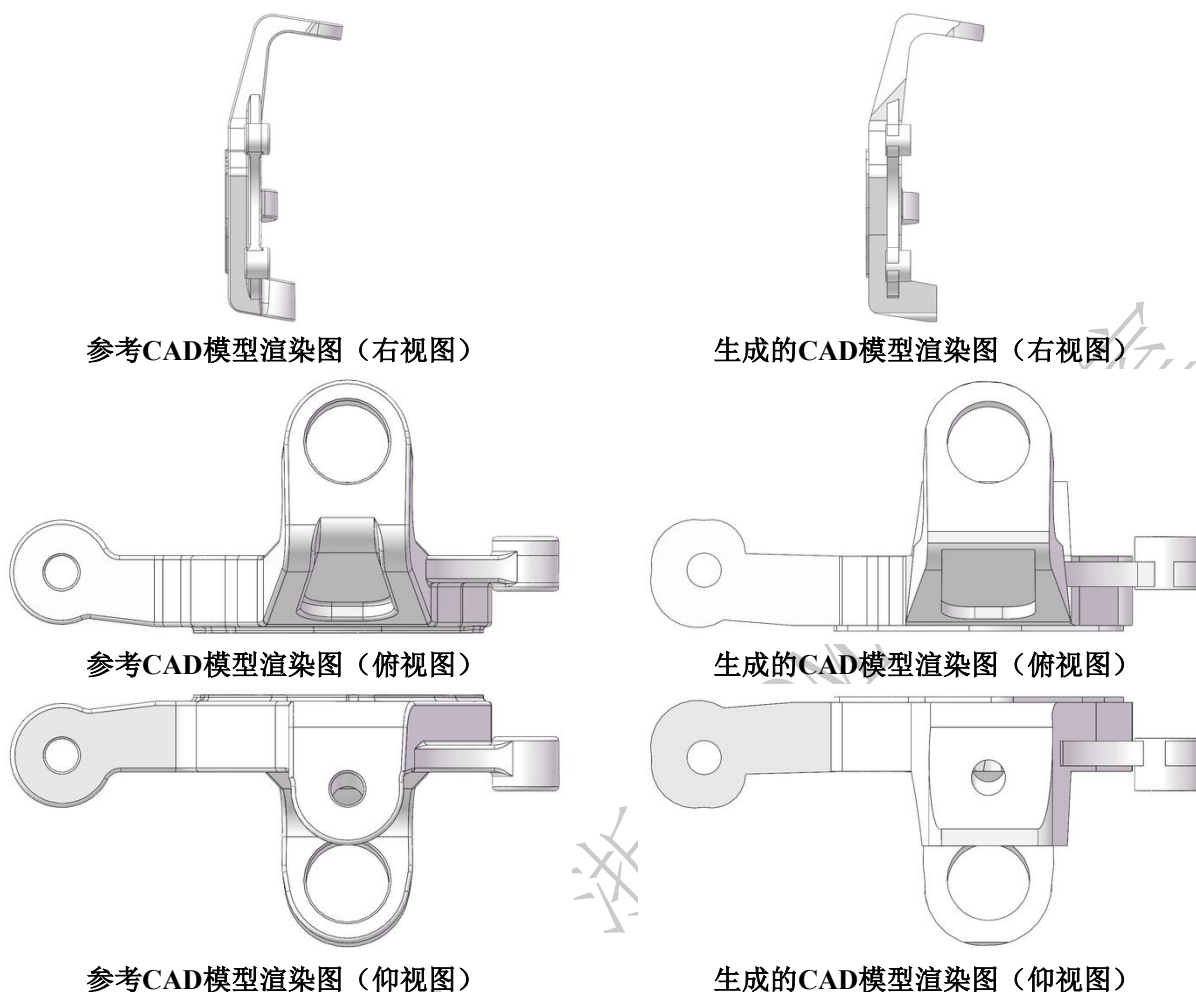


图6-12 I2C-04参考与生成模型的图

结果分析：I2C-04的生成模型实体有效，中央孔系、侧向安装耳及上部弯臂均已生成，主要特征及连接关系与参考模型基本一致。安装孔周围的台阶、底部轮廓和边缘圆角仍有简化。生成模型包围盒最大边长为237.03 mm，参考模型为317.22 mm，整体尺寸偏小，HD为45.211 mm，孔位、壁厚及安装尺寸仍有偏差。

各视图的轮廓IoU随迭代次数的变化见图6-13。



图6-13 I2C-04的GPT内部迭代次数与IoU曲线

I2C-05 GE jet engine bracket IVO (复杂)

输入渲染图

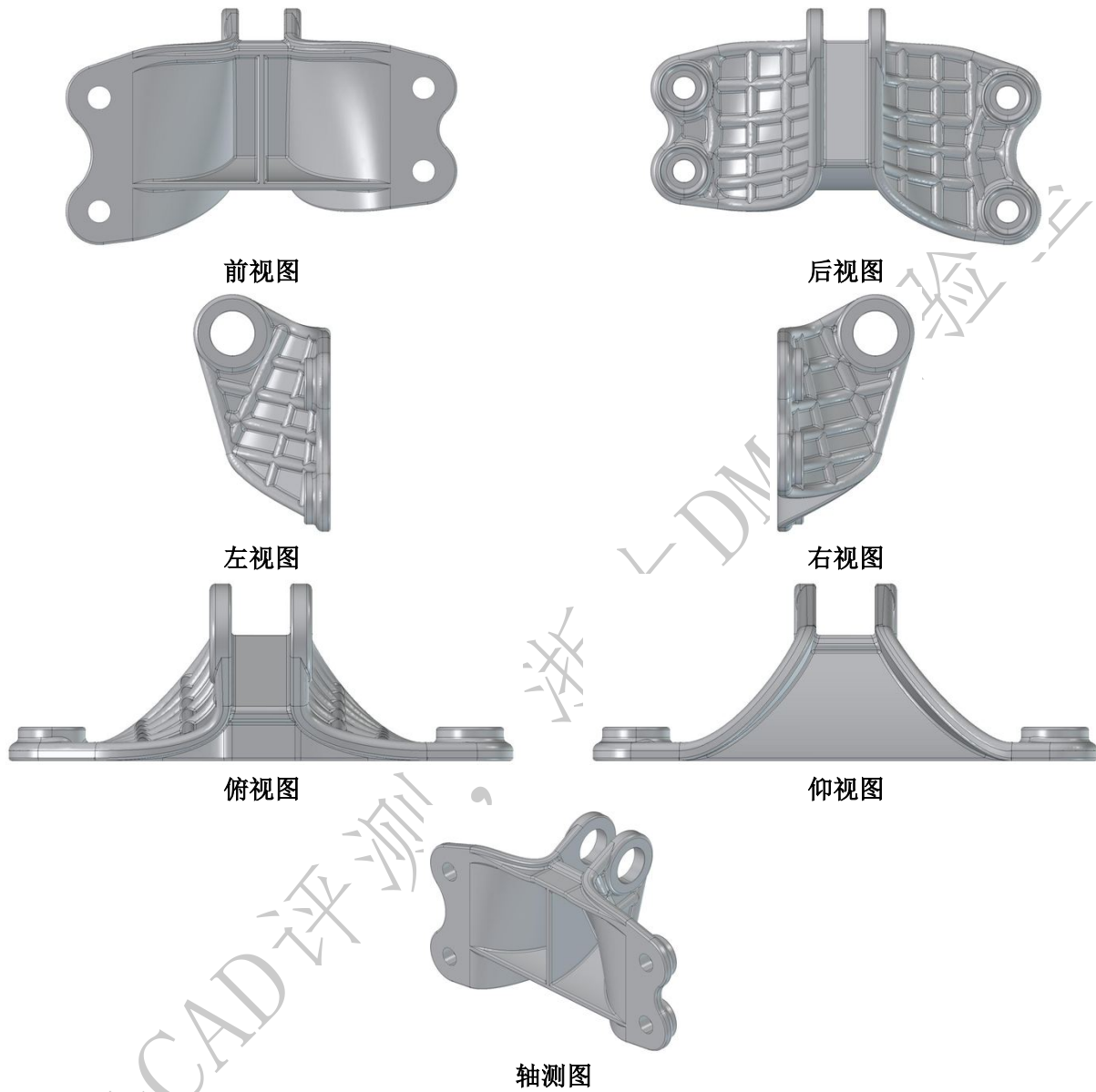
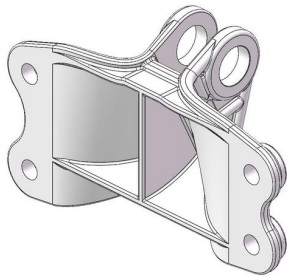
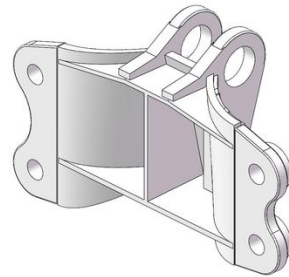


图6-14 I2C-05的七张输入渲染图

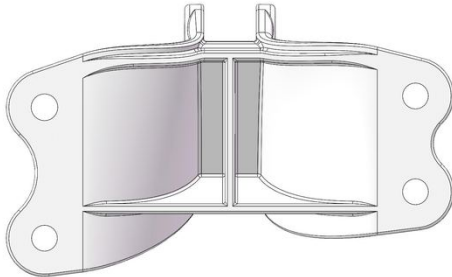
结果展示



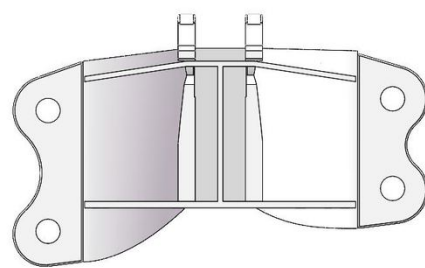
参考CAD模型渲染图（轴测图）



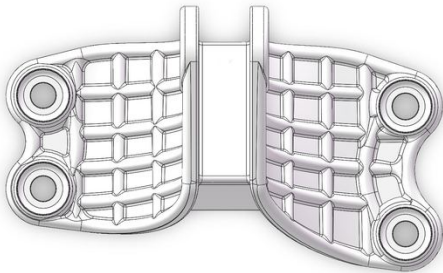
生成的CAD模型渲染图（轴测图）



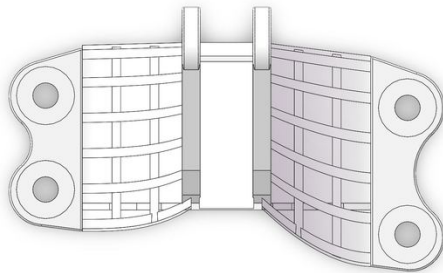
参考CAD模型渲染图（前视图）



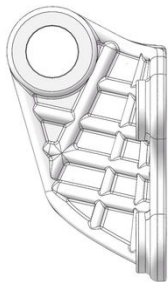
生成的CAD模型渲染图（前视图）



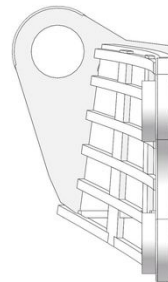
参考CAD模型渲染图（后视图）



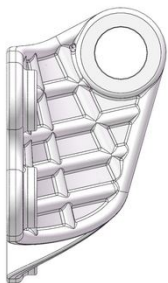
生成的CAD模型渲染图（后视图）



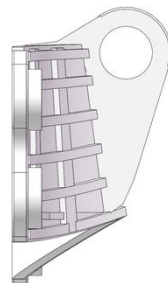
参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）



参考CAD模型渲染图（右视图）



生成的CAD模型渲染图（右视图）

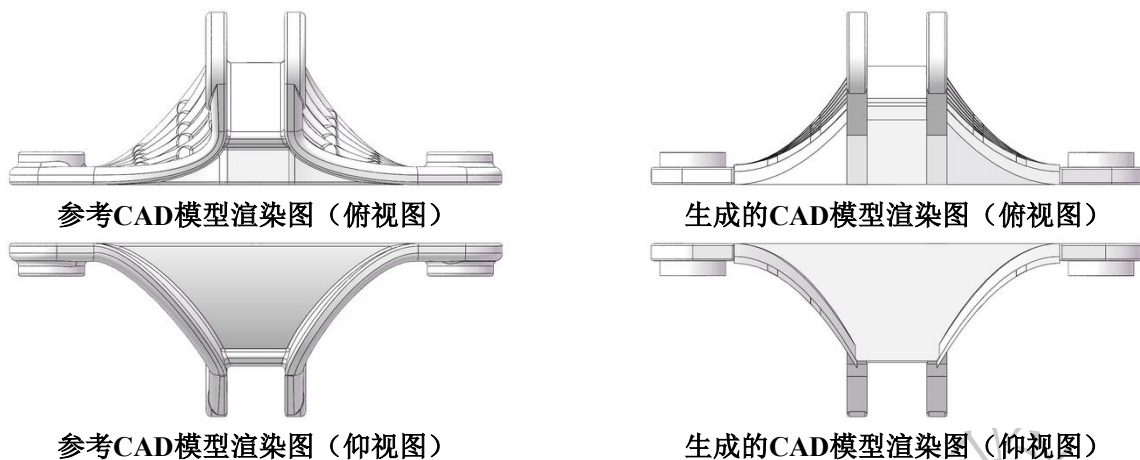


图6-15 I2C-05参考与生成模型的渲染图

结果分析：I2C-05的生成模型实体有效，非对称安装耳、双耳孔及主要支承结构已生成。翼面与耳座的连接过渡不够光滑，背面加强筋被简化为矩形截面的条带，筋板交汇处的筋板根部的过渡圆角未完整生成，细节如下图所示。HD为9.959 mm，生成模型包围盒最大边长约为160.07 mm，参考模型为178.58 mm，局部曲面和整体尺寸均有偏差。

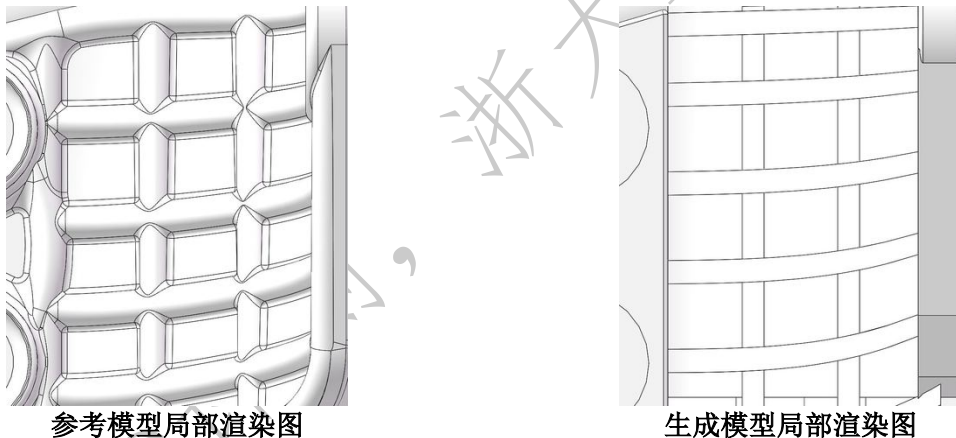


图6-16 I2C-05背面加强筋及圆角对比

各视图的轮廓IoU随迭代次数的变化见图6-17。

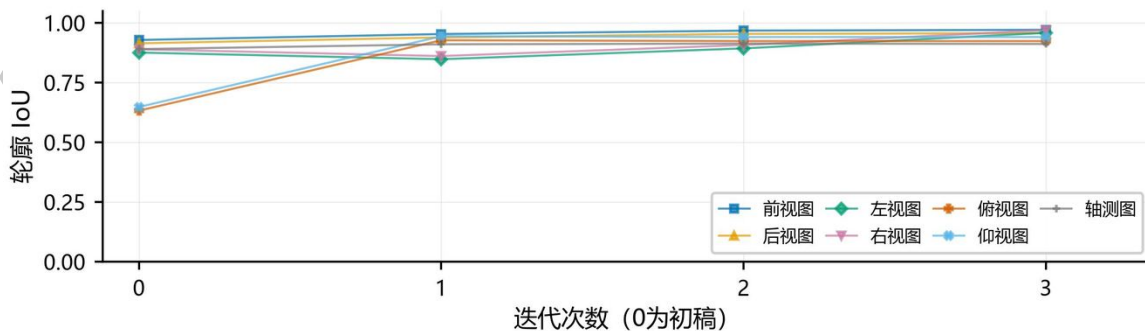


图6-17 I2C-05的GPT内部迭代次数与IoU曲线

I2C-06 Hammerhead机器人底盘（复杂）

输入渲染图

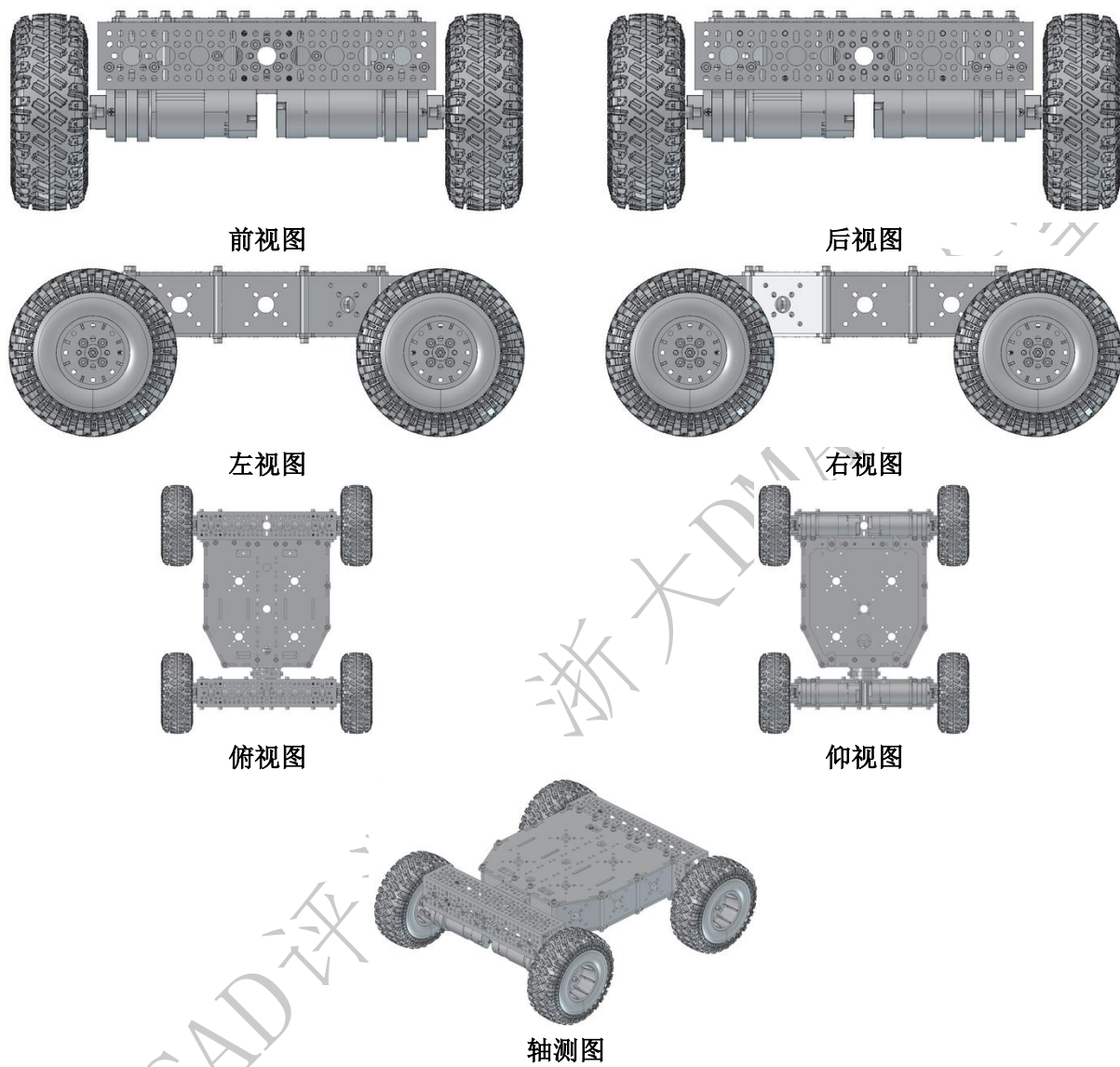
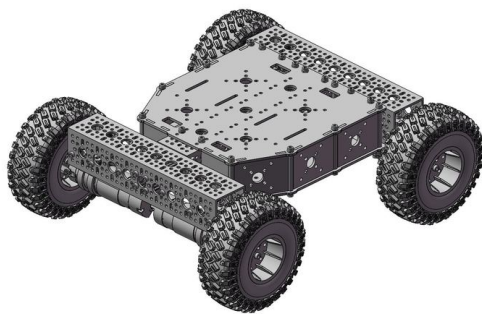
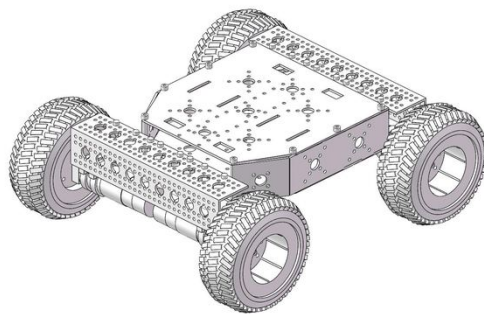


图6-18 I2C-06的七张输入渲染图

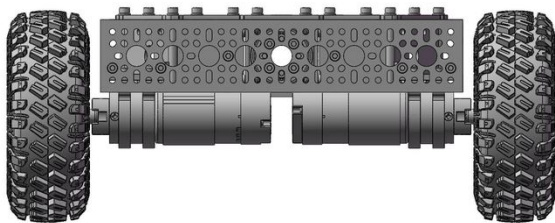
结果展示



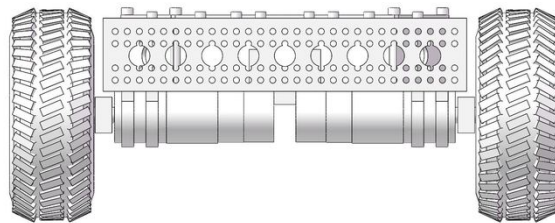
参考CAD模型渲染图（轴测图）



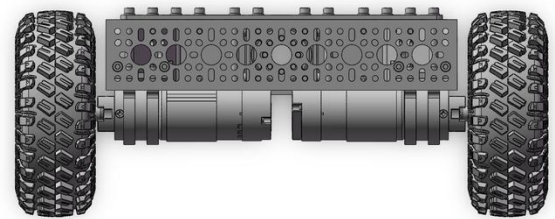
生成的CAD模型渲染图（轴测图）



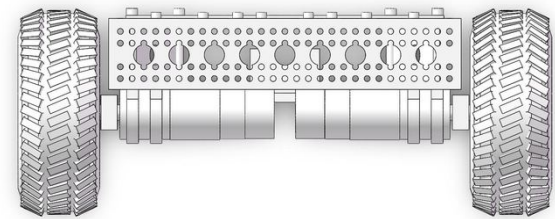
参考CAD模型渲染图（前视图）



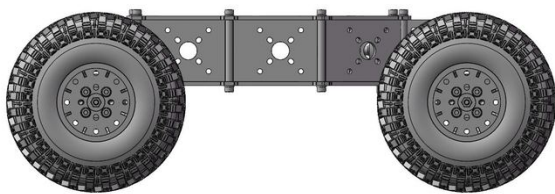
生成的CAD模型渲染图（前视图）



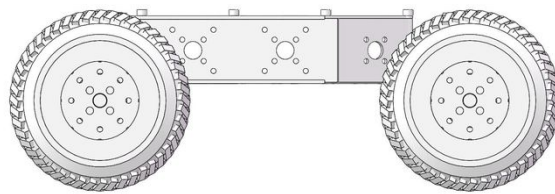
参考CAD模型渲染图（后视图）



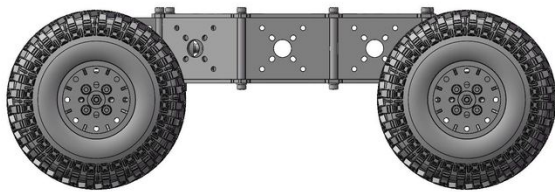
生成的CAD模型渲染图（后视图）



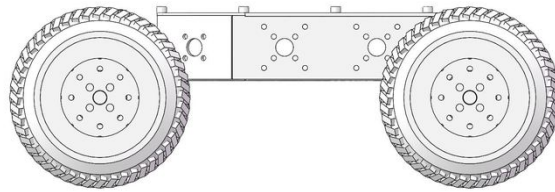
参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）



参考CAD模型渲染图（右视图）



生成的CAD模型渲染图（右视图）

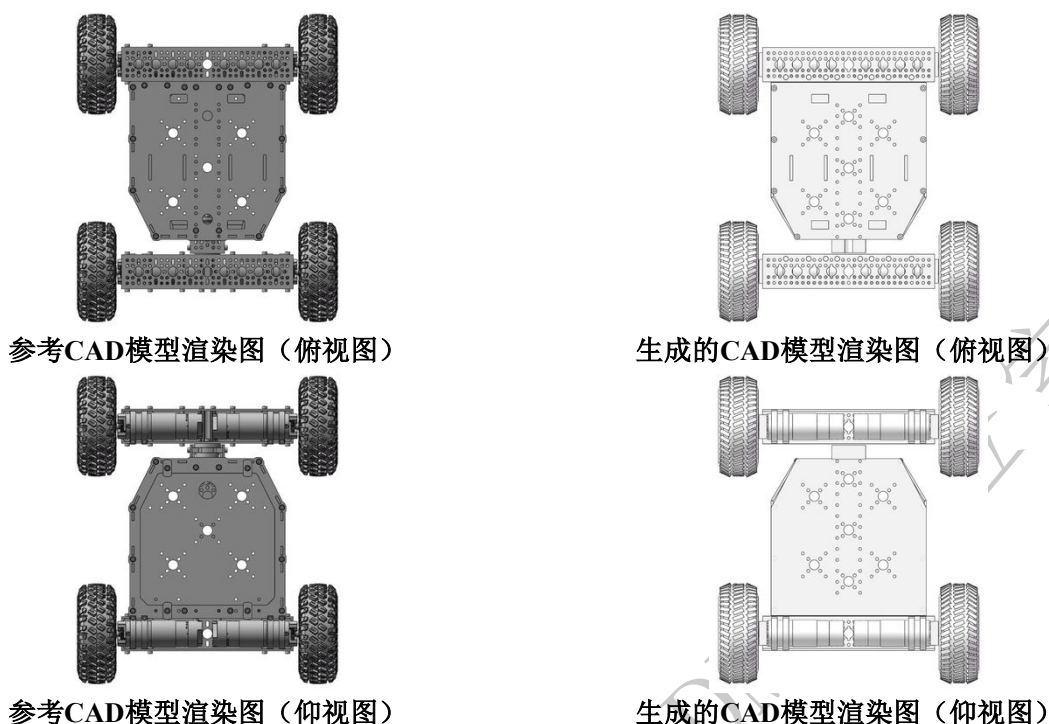


图6-19 I2C-06参考与生成模型的渲染图（续）

结果分析：I2C-06的生成模型包含721个实体，各实体通过基本有效性检查，四轮、板件及前后横梁的总体布置与参考模型接近。轮毂、板面孔阵、支承位置和连接细节仍有偏差，密集部件之间的对应关系也未逐项验证。生成底盘的最大方向跨度约为308.47 mm，参考模型为429.45 mm，整体偏小，HD为80.297 mm，恢复了底盘的主要组成，但实体数量不能直接作为装配零件数。

各视图的轮廓IoU随迭代次数的变化见图6-20。

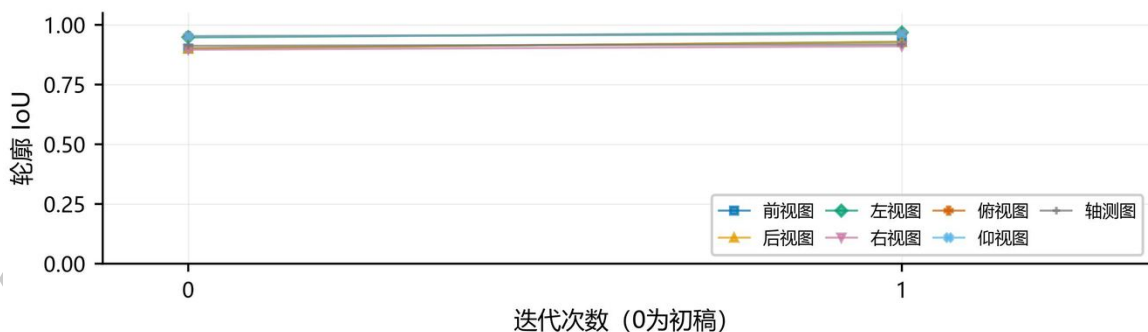


图6-20 I2C-06的GPT内部迭代次数与IoU曲线

6.3.2 扩展实验结果

以下分别展示三视图、六视图、关闭网络搜索和复杂壳体压力测试的结果。S1、S2以相同样本的七视图主实验为对照，重点比较视图减少后主体结构、局部特征和三维尺寸的

变化。HD的计算均采用同一参考STEP和相同计算方法。各用例展示一张参考模型轴测图和一张生成模型轴测图。

(1) S1视图减少

本组仅提供前、右、俯三个正投影视图，网络搜索保持开启，Prompt与主实验一致。结果汇总见表6-7。

表6-7 S1三视图扩展实验结果概览

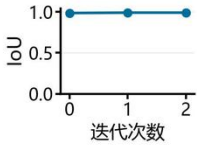
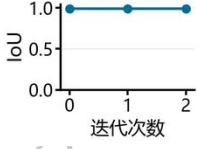
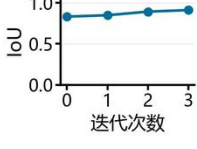
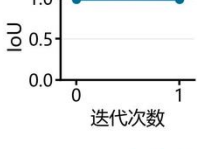
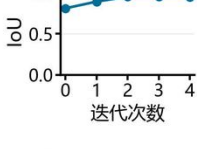
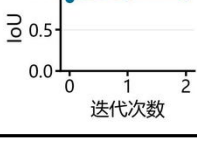
样本ID	实体有效性	三视图平均IoU	HD距离 (mm)	IoU与迭代次数曲线图	对照观察的主要差异
I2C-01	通过	0.9912	15.147		卡扣中部及角部孔座有差异，整体尺度偏大。
I2C-02	通过	0.9910	27.924		开口下沿错误连通，整体尺寸大于参考模型。
I2C-03	通过	0.9123	112.981		衬套与主体分为3个实体，连接简化且尺度明显偏小。
I2C-04	通过	0.9821	21.494		连接壁呈分段折面；外形尺寸接近，局部过渡仍简化。
I2C-05	通过	0.9514	14.098		前侧筋条方向与参考不同，翼面过渡简化。
I2C-06	通过	0.9378	78.811		底盘布局保留；尺度偏小，孔阵及连接细节存在差异。

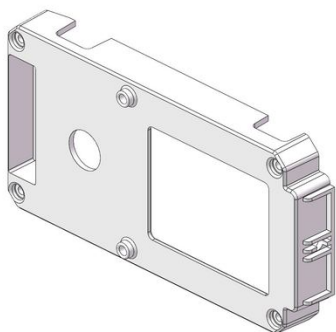
表6-8 图像到CAD扩展S1实验各样本测试用时

样本ID	I2C-01	I2C-02	I2C-03	I2C-04	I2C-05	I2C-06
用时	13分钟	13分钟	19分钟	10分钟	24分钟	20分钟

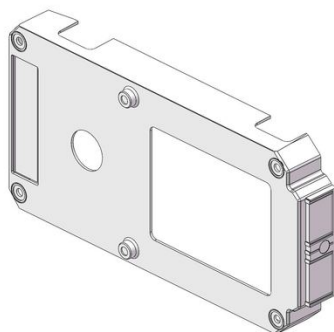
总体对比分析：三视图组的六个样本均通过单个实体的有效性检查，主体轮廓基本保留。但I2C-03由多个独立实体组成，衬套与主体存在空间交叠，未形成参考零件的单实体连接。

差异主要出现在侧面卡扣、孔座连接和曲面筋条等局部结构。HD相对主实验：I2C-03由62.401 mm增至112.981 mm，而I2C-02、I2C-04分别降至27.924 mm和21.494 mm，主要与后两例的生成尺寸更接近参考模型有关，局部曲面和连接特征仍有偏差。

I2C-01 TrackerOne背板（简单）



参考CAD模型渲染图

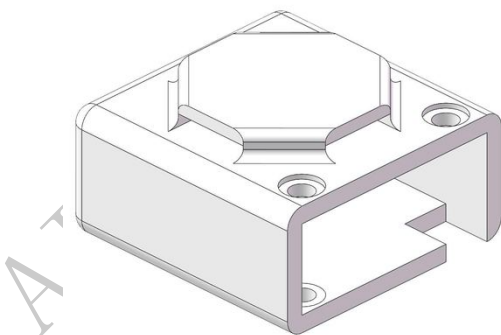


S1生成的CAD模型渲染图

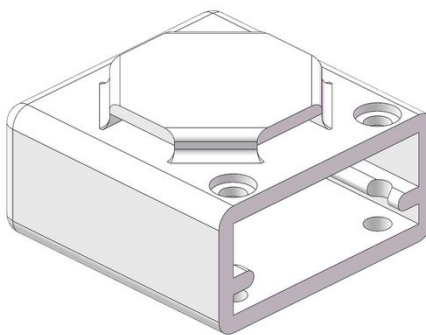
图6-21 I2C-01三视图输入生成CAD对比结果

对比分析：三视图下，背板的窗口、圆孔和周边翻边仍能恢复。侧面卡扣被改为带圆形凹部的分段结构，角部孔座及小过渡与参考模型不同。其HD为15.147 mm，与主实验的14.758 mm接近；两组沿长边方向均约为158 mm，大于参考模型的132.11 mm。减少视图后主体布局变化较小，卡扣等局部结构的差别更明显。

I2C-02 KnowCO2传感器外壳（简单）



参考CAD模型渲染图

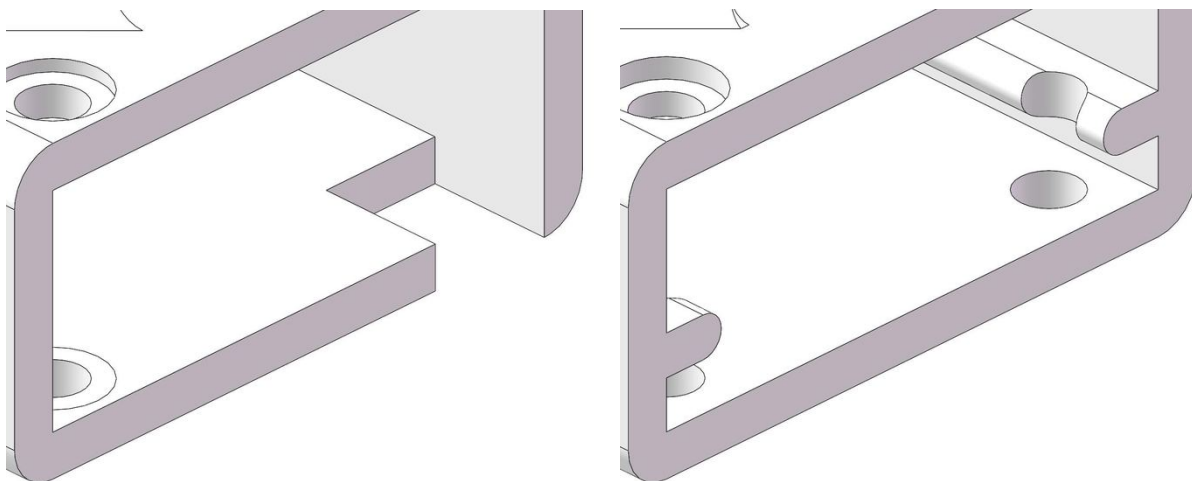


S1生成的CAD模型渲染图

图6-22 I2C-02三视图输入生成CAD对比结果

对比分析：三视图下，壳体和顶部平台的外形得到保留，但侧向开口下沿被连成完整边框，与参考模型的开边缺口不同，内部导向结构也未准确对应。生成模型尺寸为64.00×62.25×36.00 mm，大于参考模型，但比主实验更接近参考尺寸，HD降至27.924 mm。

距离指标降低主要伴随尺寸偏差减小；连续下框占用了参考模型预留的开口空间，可能影响配合件装入和安装。

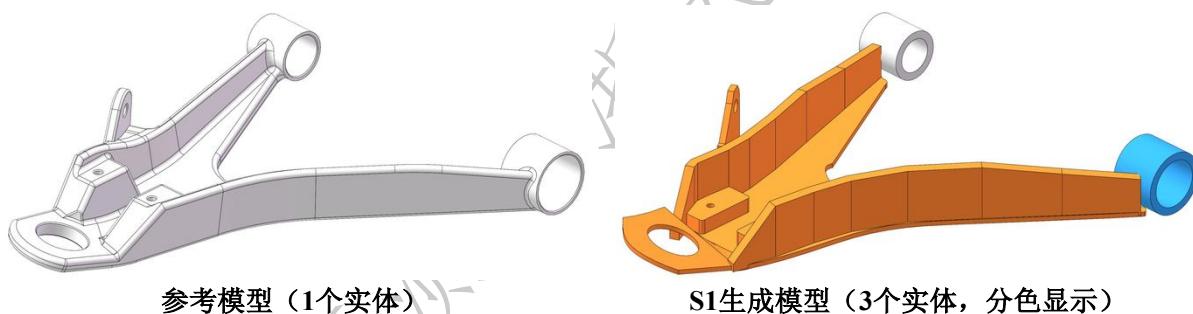


参考模型局部渲染图

S1生成模型局部渲染图

图6-23 I2C-02三视图结果的开口边界对比

I2C-03 C5 Corvette下控制臂（中等）

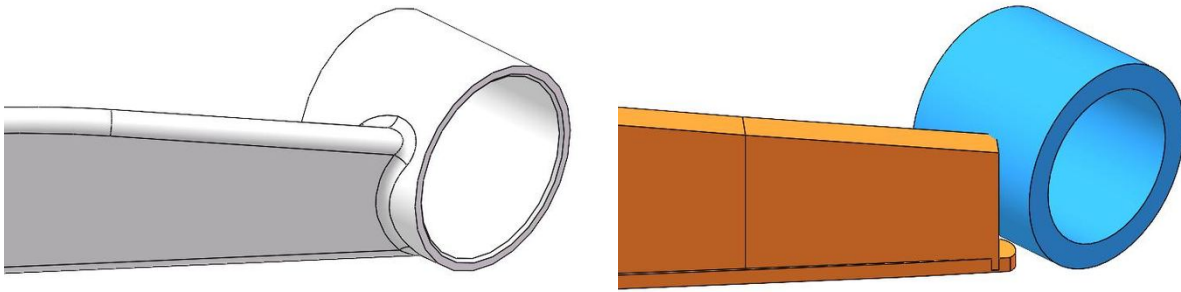


参考模型（1个实体）

S1生成模型（3个实体，分色显示）

图6-24 I2C-03三视图输入生成CAD对比结果

对比分析：三视图下，叉臂、衬套及前端耳板均已生成，但参考模型的1个实体被建成3个独立实体，图中以不同颜色区分。进一步对实体两两求交，两个衬套与主体的交叠体积分别约为 0.454 mm^3 和 135.024 mm^3 ，说明连接处存在实体重叠，且未完成布尔合并。局部截图可见衬套根部的过渡圆角也未生成。各实体虽有效，单零件的整体拓扑仍与参考模型不一致。生成模型包围盒最大边长为 315.52 mm ，小于参考模型的 488.73 mm ，HD为 112.981 mm 。

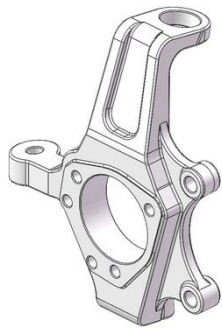


参考模型：单实体过渡

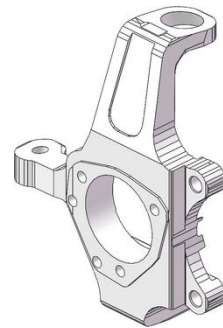
S1生成模型：实体未合并

图6-25 I2C-03三视图结果的衬套连接对比

I2C-04 C5 Corvette转向节（中等）



参考CAD模型渲染图

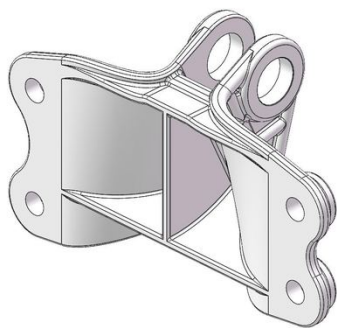


S1生成的CAD模型渲染图

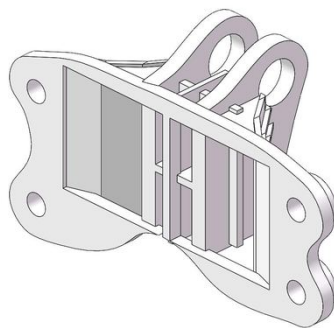
图6-26 I2C-04三视图输入生成CAD对比结果

对比分析：三视图下，中央孔系、上部弯臂和侧向安装耳均已生成。图中连接壁和安装耳出现分段折面，孔口台阶及铸造圆角没有完整恢复。HD为21.494 mm，低于主实验的45.211 mm；本组最大方向跨度为315.83 mm，接近参考模型的317.22 mm，模型整体尺寸更接近，误差集中在孔口台阶、耳板连接及圆角过渡，相关接触面仍与参考零件不同。

I2C-05 GE jet engine bracket IVO (复杂)



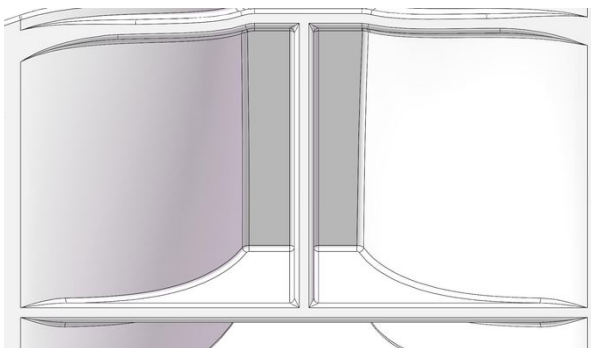
参考CAD模型渲染图



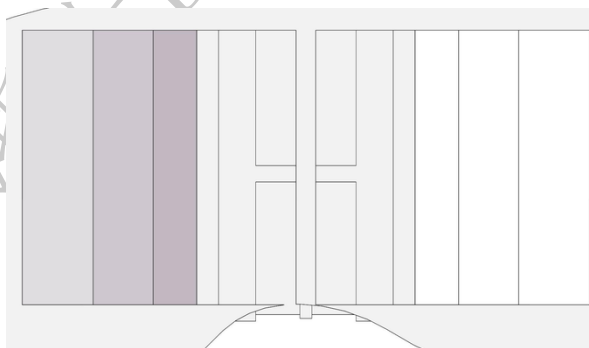
S1生成的CAD模型渲染图

图6-27 I2C-05三视图输入生成CAD对比结果

对比分析：三视图下，非对称安装耳和双耳孔的位置基本正确，但前侧横向加强筋被改为多道竖向直筋，翼面也出现分段，局部对比如下图所示。HD为14.098 mm，高于主实验的9.959 mm，整体尺寸仍偏小。参考模型底部原本为复杂自由曲面，但是生成模型却为简单面堆叠而成。



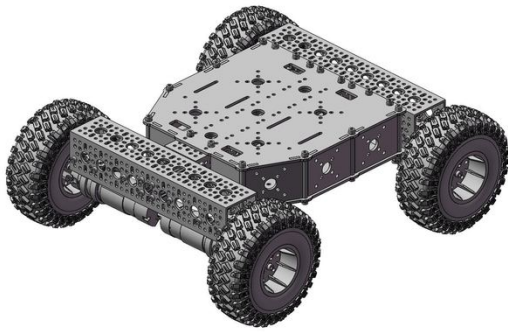
参考模型局部渲染图



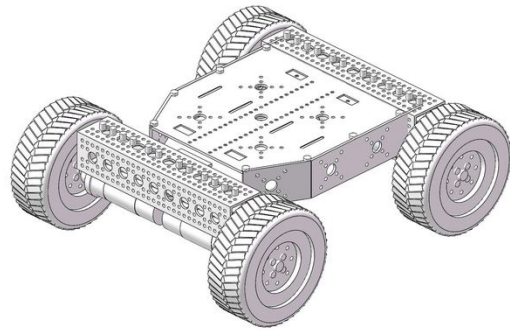
S1生成模型局部渲染图

图6-28 I2C-05三视图结果的加强筋方向对比

I2C-06 Hammerhead机器人底盘（复杂）



参考CAD模型渲染图



S1生成的CAD模型渲染图

图6-29 I2C-06三视图输入生成CAD对比结果

对比分析：三视图下，底盘板件、横梁和四轮的总体布置得到保留。输出包含480个有效实体，与主实验的721个不同，但这一差别不能直接解释为遗漏了241个零件，还涉及轮胎纹路等细节采用分体或合并表达的区别。该组HD为78.811 mm，与主实验的80.297 mm接近，整体尺度仍明显偏小；孔阵、轮毂及连接结构未逐一对应。

(2) S2移除轴测图

本组提供前、后、左、右、俯、仰六个正投影视图，移除轴测图，网络搜索和Prompt与主实验保持一致。结果汇总见表6-9。

表6-9 S2六视图扩展实验结果概览

样本ID	实体有效性	六视图平均IoU	HD距离 (mm)	IoU与迭代次数曲线图	对照观察的主要差异
I2C-01	通过	0.9896	17.246		卡扣简化为平直槽形，孔座细节及尺度仍有偏差。
I2C-02	通过	0.9947	36.380		开口下沿错误连通，尺寸介于三视图组和主实验之间。
I2C-03	通过	0.9467	63.106		形成单实体；尺度接近主实验，眼板与筋壁过渡仍简化。
I2C-04	通过	0.9806	20.867		外形尺寸接近参考，孔口及连接处圆角仍有简化。

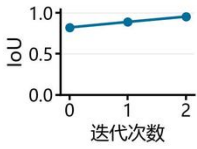
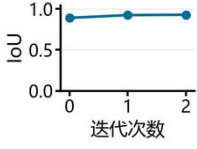
样本ID	实体有效性	六视图平均IoU	HD距离 (mm)	IoU与迭代次数曲线图	对照观察的主要差异
I2C-05	通过	0.9549	14.616		正面竖向直筋与参考布局不同，曲面过渡未准确恢复。
I2C-06	通过	0.9291	72.731		总体布局保留，尺寸较主实验接近参考，连接细节仍有偏差。

表6-10 图像到CAD扩展S2实验各样本测试用时

样本ID	I2C-01	I2C-02	I2C-03	I2C-04	I2C-05	I2C-06
用时	9分钟	11分钟	17分钟	17分钟	20分钟	17分钟

总体对比分析：移除轴测图后，六个样本均通过基本实体检查，主体布局总体保持。I2C-03的HD由三视图组的112.981 mm降至63.106 mm，已接近七视图主实验的62.401 mm；I2C-05在六视图下仍未正确恢复正面筋条布局。I2C-02、I2C-04的HD低于主实验，但主要伴随整体尺寸更接近参考模型。

I2C-01 TrackerOne背板（简单）

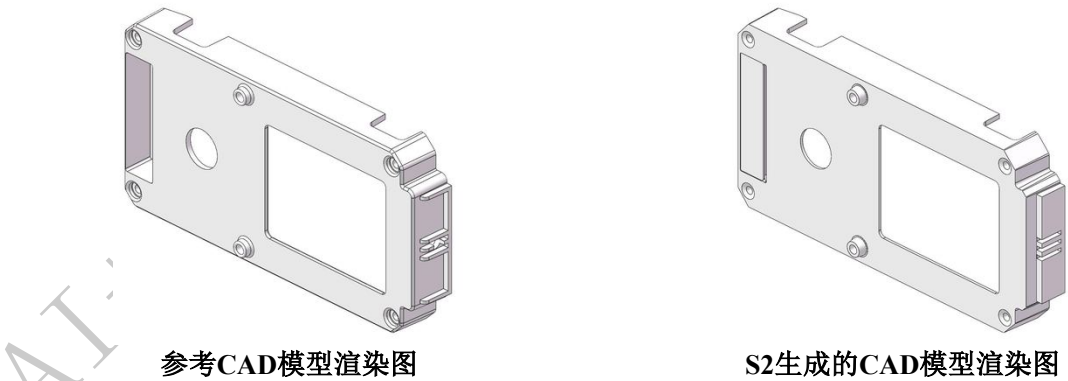
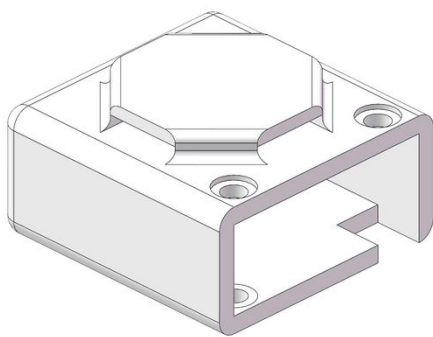


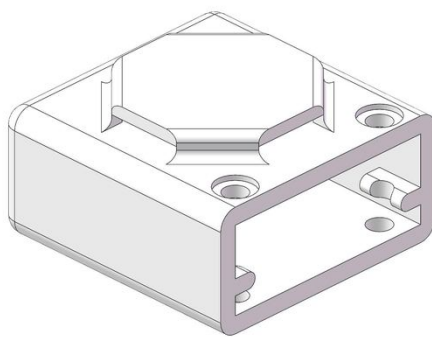
图6-30 I2C-01六视图输入生成CAD对比结果

对比分析：六视图下，窗口和孔位的总体布局保持，侧面卡扣仍被简化为平直槽形，角部孔座的环形细节与参考模型存在差异。其HD为17.246 mm，略高于三视图组及主实验。生成模型沿长边方向约为159.90 mm，大于参考模型的132.11 mm，生成模型局部卡扣形状未得到准确恢复。

I2C-02 KnowCO2传感器外壳（简单）



参考CAD模型渲染图

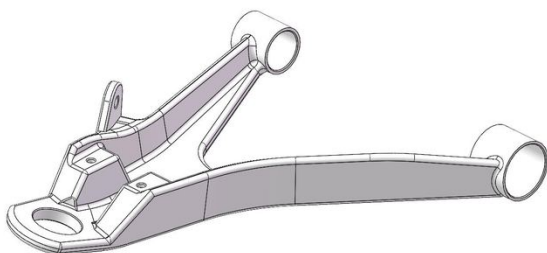


S2生成的CAD模型渲染图

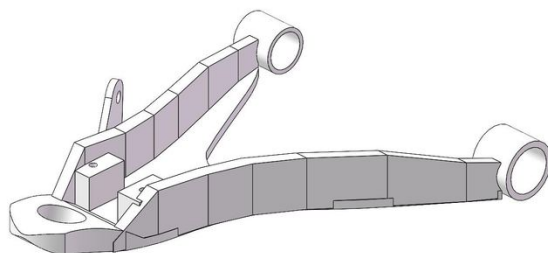
图6-31 I2C-02六视图输入生成CAD对比结果

对比分析：六视图下，顶部平台及根部过渡与参考模型较接近，但侧向开口仍保留连续下沿，与三视图组存在相同的结构偏差。生成模型尺寸为75.00×73.00×42.00 mm，介于三视图组和主实验之间，HD为36.380 mm，轮廓IoU为0.9947，但生成模型尺寸明显放大，开口下缘的连续边框还改变了原有避让空间。这些偏差涉及配合件的定位与装入，需进一步核验。

I2C-03 C5 Corvette下控制臂（中等）



参考CAD模型渲染图

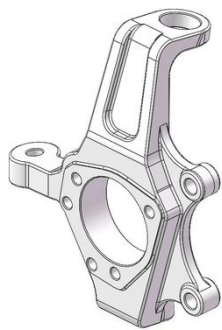


S2生成的CAD模型渲染图

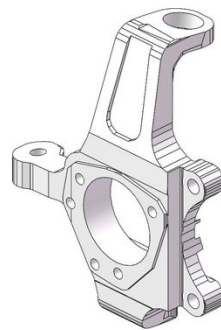
图6-32 I2C-03六视图输入生成CAD对比结果

对比分析：六视图下，输出形成1个有效实体，叉臂、衬套及前端眼板的连接较三视图组完整。生成模型最大方向跨度增至394.50 mm，HD由三视图组的112.981 mm降至63.106 mm，与主实验接近。但眼板根部及孔口仍以平面、分段曲面近似原有自由曲面和圆角过渡。虽然已合为单实体，局部连接截面和过渡形状仍与参考零件不同。

I2C-04 C5 Corvette转向节（中等）



参考CAD模型渲染图

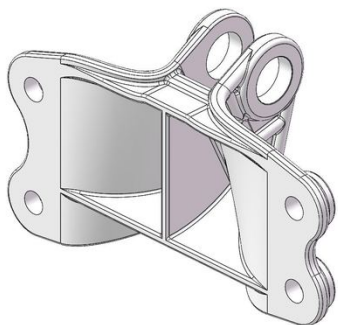


S2生成的CAD模型渲染图

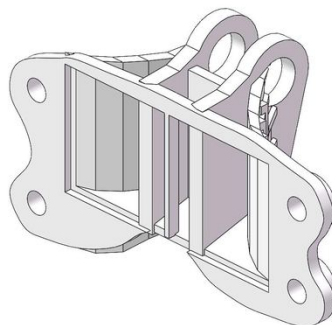
图6-33 I2C-04六视图输入生成CAD对比结果

对比分析：六视图下，主要孔系、弯臂和安装耳得到保留，最大方向跨度为316.00 mm，与参考模型的317.22 mm接近。HD为20.867 mm，与三视图组相近，低于主实验。安装耳和连接壁的曲面过渡仍被分段平面近似，孔口台阶及圆角也有缺失。偏差集中在局部接触面、壁厚和连接过渡，涉及耳板的装配配合及局部受力。

I2C-05 GE jet engine bracket IVO（复杂）



参考CAD模型渲染图

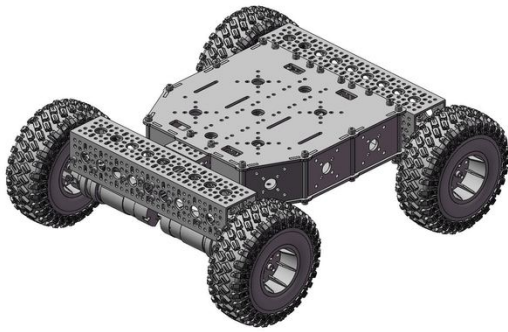


S2生成的CAD模型渲染图

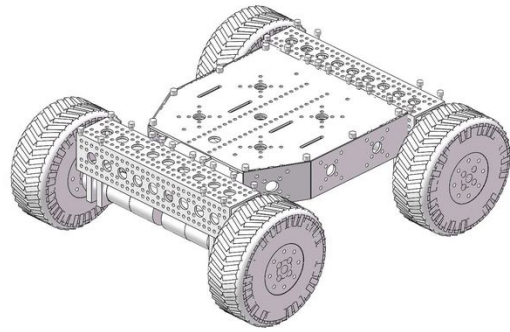
图6-34 I2C-05六视图输入生成CAD对比结果

对比分析：六视图下，翼面两端的安装孔和顶部双耳孔得到保留，但正面仍布置多道竖向直筋，与参考模型的横向筋条布局不同。其HD为14.616 mm，与三视图组接近，高于主实验的9.959 mm。该组与主实验的外形尺寸相近，筋条的数量、方向和连接方式却有明显差异，薄翼的局部支承方式未按参考结构恢复。

I2C-06 Hammerhead机器人底盘（复杂）



参考CAD模型渲染图



S2生成的CAD模型渲染图

图6-35 I2C-06六视图输入生成CAD对比结果

对比分析：六视图下，板件、横梁和四轮的总体布置基本还原，输出包含799个有效实体。包围盒最大边长为317.33 mm，比三视图组和主实验更接近参考模型的429.45 mm，HD降至72.731 mm，但整体尺寸仍偏小。轮毂、孔阵列及内部连接特征存在差异，实体有效性检查尚不能代替接口尺寸和装配配合检查。

(3) S3关闭网络搜索

本组对I2C-05和I2C-06关闭外部网络检索，仍提供六视图加轴测图，并沿用主实验Prompt，结果汇总见表6-11。

表6-11 S3关闭网络搜索扩展实验结果概览

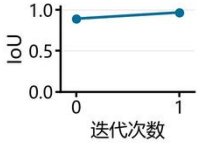
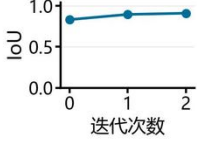
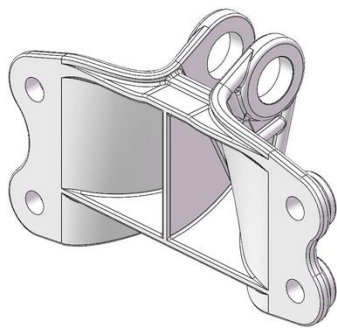
样本ID	实体有效性	六视图加轴测图 平均IoU	HD距离 (mm)	IoU与迭代次数曲线图	对照观察的主要差异
I2C-05	通过	0.9690	19.905		前侧以连续板面近似，翼面及主要筋条未完整恢复。
I2C-06	通过	0.9109	146.769		轮胎、轮毂及密集孔阵明显简化，整体尺度偏小。

表6-12 图像到CAD扩展S3实验各样本测试用时

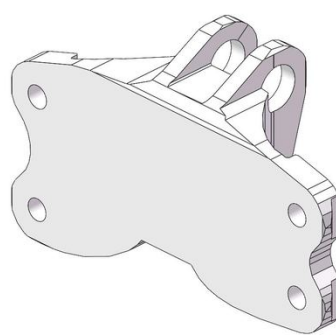
样本ID	I2C-05	I2C-06
用时	24分钟	20分钟

总体对比分析：两例生成了通过实体有效性检查的模型。I2C-05经1次修改，七视图平均IoU由0.8927升至0.9690，但HD为19.905 mm，高于联网主实验的9.959 mm；I2C-06经2次修改，平均IoU由0.8325升至0.9109，HD为146.769 mm，也高于主实验的80.297 mm。两例均保留了主体轮廓，但支架的曲面筋条、底盘的轮胎和孔阵细节明显简化。

I2C-05 GE jet engine bracket IVO（复杂）



参考CAD模型渲染图

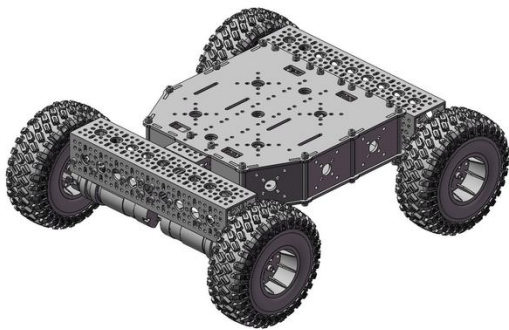


S3生成的CAD模型渲染图

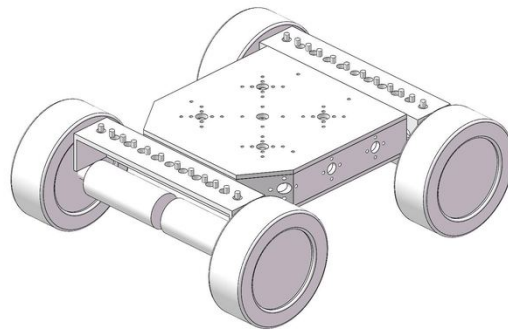
图6-36 I2C-05无网络搜索生成CAD对比结果

对比分析：生成模型保留了两侧安装孔和顶部双耳孔，但前侧以连续平板替代原有自由曲面及筋板结构，弯曲翼面、横向加强筋和局部圆角未完整恢复。其外形尺寸与联网主实验接近，HD却由9.959 mm增至19.905 mm，IoU为0.9690。平板近似改变了翼面的曲率、局部截面及内部留空，支承几何与参考模型不同；较高的轮廓重合程度不足以说明原有结构功能得到保留。

I2C-06 Hammerhead机器人底盘（复杂）



参考CAD模型渲染图

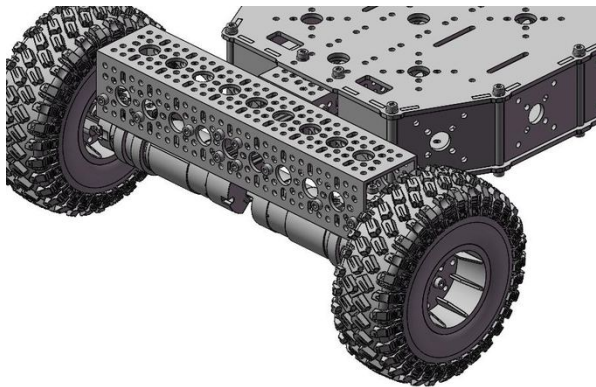


S3生成的CAD模型渲染图

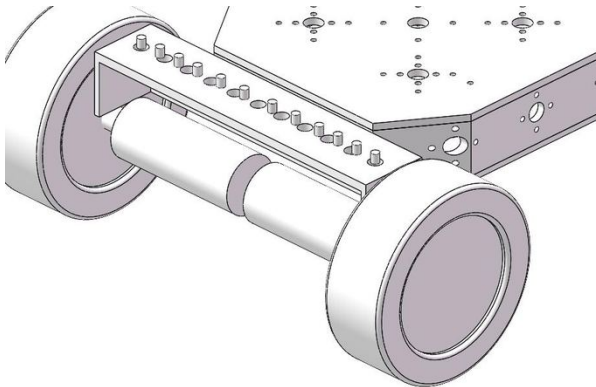
图6-37 I2C-06无网络搜索生成CAD对比结果

对比分析：底板、顶板、横梁和四轮的总体布局得到保留，输出包含45个有效实体。轮胎表面没有恢复胎纹，轮毂采用较简单的圆盘结构，横梁和板面的密集孔槽也有明显缺失。最大方向跨度约为201.58 mm，小于参考模型的429.45 mm，也小于联网主实验的

308.47 mm；HD增至146.769 mm。除整体尺度偏小外，轮毂和安装孔槽的简化还改变了构件的连接几何及装入位置，后续需核验装配和运动关系。



参考模型局部渲染图



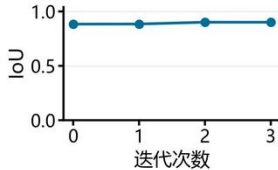
S3生成模型局部渲染图

图6-38 I2C-06无网络结果的车轮及横梁特征对比

(4) S4压力测试

压力测试选取I2C-07钟形壳体，包含多层法兰、孔座、侧窗及密集筋条，输入为六视图加轴测图，网络搜索保持开启，用于进一步检查复杂样本GPT-6的恢复情况，结果见表6-13。

表6-13 S4钟形壳体压力测试结果

样本ID	难度	实体有效性	六视图加轴测图 平均IoU	HD距离 (mm)	IoU与迭代次数曲线图	用时
I2C-07	复杂	通过	0.9013	78.484		20分钟

I2C-07 钟形壳体（复杂）

输入渲染图

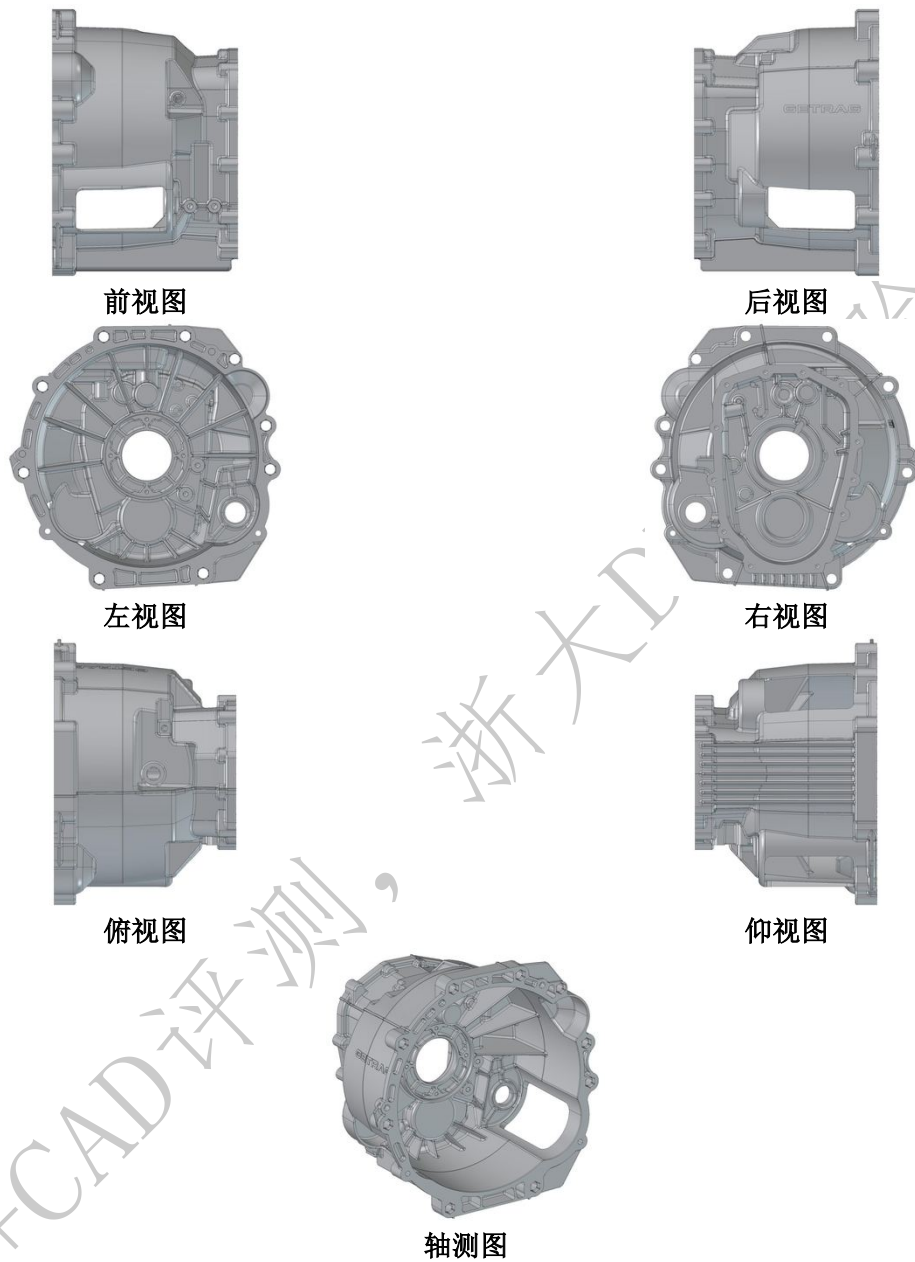
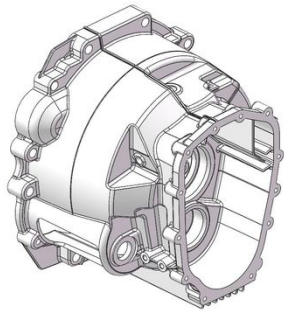
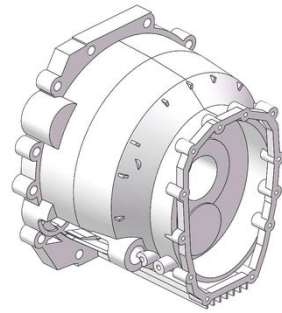


图6-39 七张输入渲染图

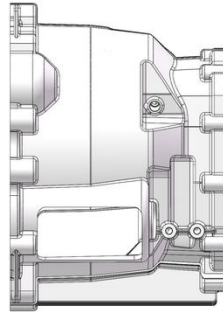
结果展示



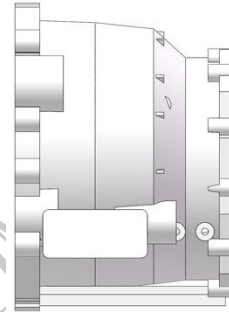
参考CAD模型渲染图（轴测图）



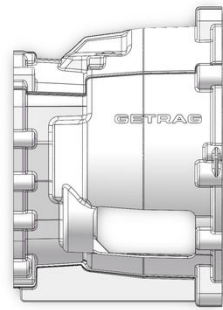
生成的CAD模型渲染图（轴测图）



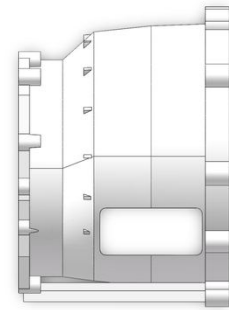
参考CAD模型渲染图（前视图）



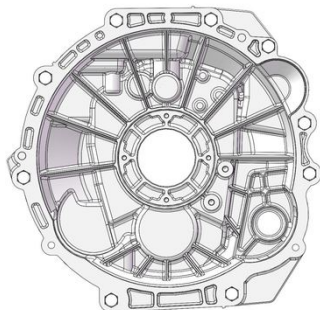
生成的CAD模型渲染图（前视图）



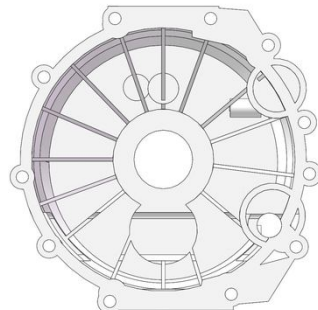
参考CAD模型渲染图（后视图）



生成的CAD模型渲染图（后视图）



参考CAD模型渲染图（左视图）



生成的CAD模型渲染图（左视图）

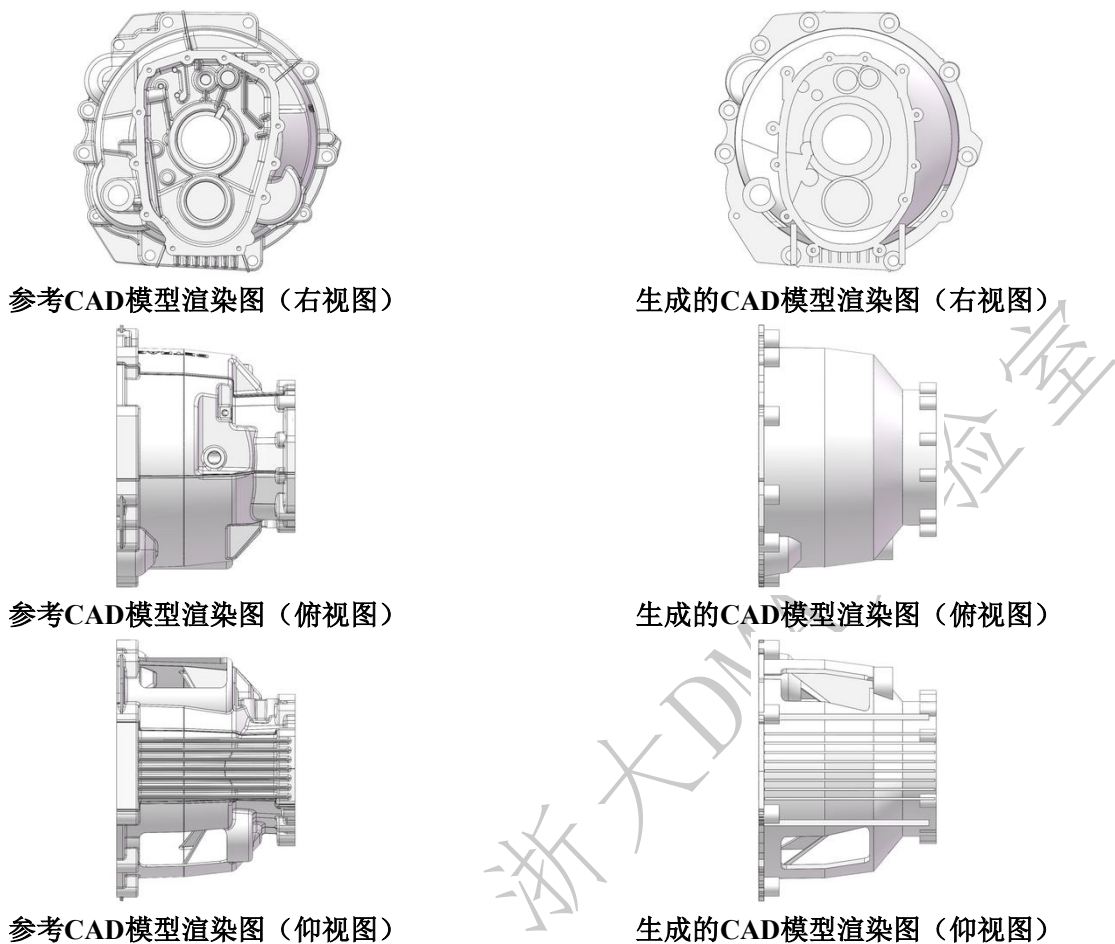


图6-40 I2C-07钟形壳体生成CAD对比结果

结果分析：钟形壳体的生成结果包含1个有效实体，大端法兰、主轴孔、侧窗和径向筋等主要结构已经生成。与参考模型相比，外壁和法兰之间的连接被明显简化，孔座周围的铸造过渡、细筋及局部凹槽未完整恢复，局部自由曲面及连接过渡被规则孔座和平直筋板近似。其七视图平均IoU为0.9013，HD为78.484 mm；生成模型最大方向跨度为302.99 mm，参考模型为392.17 mm，整体尺寸明显偏小。

各视图的轮廓IoU随迭代次数的变化见图6-41。

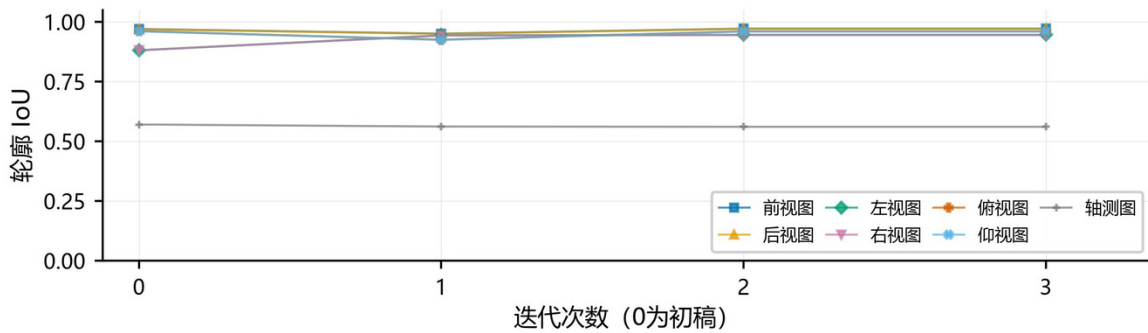
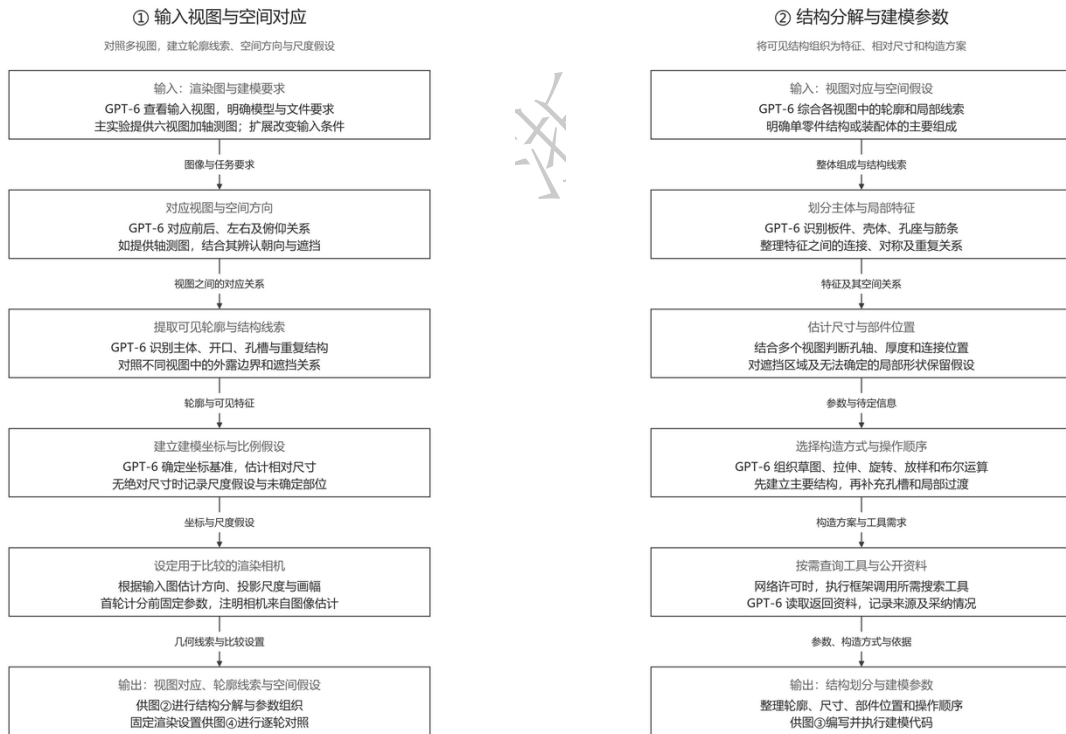


图6-41 I2C-07的GPT内部迭代次数与IoU曲线

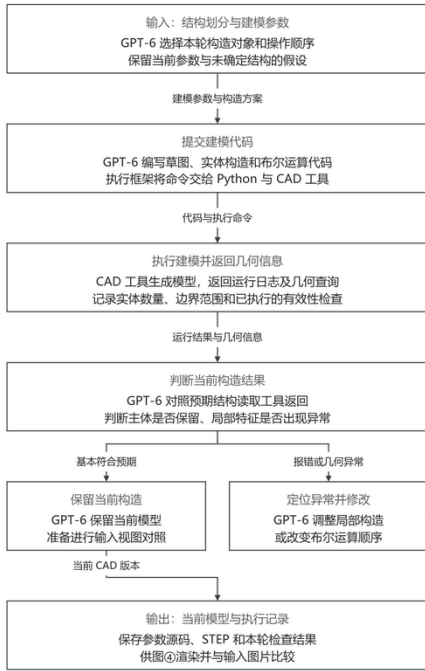
6.4 GPT-6推理过程分析

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对图像到CAD模型重建/生成任务可能采用如下推理流程：



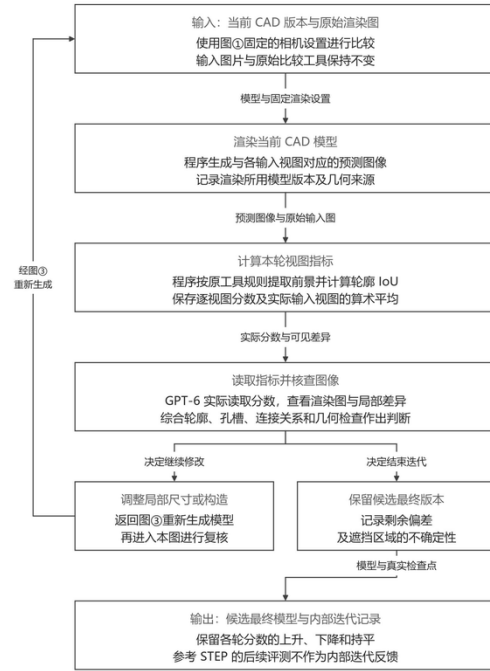
③ 实体建模与几何检查

执行建模代码，并根据返回结果处理构造异常



④ 固定视图对比与自主迭代

结合实际 IoU 与视觉核查，决定修改或结束迭代



⑤ 最终核查与结果交付

对应模型版本、检查结果和实际记录，说明验证范围



6.5 能力边界

在图像到CAD模型重建/生成任务上，GPT-6能够综合多个视图中的轮廓、孔槽和相对位置，重建零件的主体结构并生成STEP模型。本轮测试中，规则板件和壳体的主要外形通常能够恢复，复杂支架、底盘和钟形壳体也保留了主要组成及总体布局。

当前的不足主要体现在自由曲面重建、实体连接和功能相关尺寸。复杂自由曲面及筋板交汇过渡常被平面、直纹面或规则筋板近似，局部曲率、截面及圆角未准确恢复。个别单零件还由多个独立实体堆叠而成，连接处存在体积交叠且未完成布尔合并，单个实体有效不能说明整体连接正确。孔槽、卡扣、导向结构及装配接口的偏差可能影响定位、安装或承载，缺少尺寸线索时的尺度假定也会影响配合。无网络测试仍能完成模型生成，但局部结构较联网主实验更简化。

7 B-rep模型参数化重建

7.1 任务目标与评价标准

B-rep模型参数化重建任务主要测试GPT-6能否依据B-rep模型所包含的几何与拓扑信息，推断其可能的特征组成、草图与基准、建模历史及依赖关系，并重建对应的参数化CAD模型。该任务用于评价GPT-6从最终几何反推建模逻辑，并完成参数化建模规划与执行的综合能力。

考虑到同一B-rep模型可能由多种不同但合理的参数化建模历史生成，本测试不将“与原始特征树逐节点完全一致”作为唯一成功条件，而是分别评价最终几何是否正确、参数化结构是否合理、模型是否能够稳定编辑，以及重建历史与输入B-rep模型的一致程度。具体任务说明见表7-1。

表7-1 B-rep模型参数化重建测试任务说明

测试要素	说明
测试问题	GPT-6能否依据B-rep模型，恢复合理的特征组成、草图与基准、建模顺序及依赖关系，并重建一个几何一致且可编辑的参数化CAD模型？
核心输入	机械零件的B-rep模型。
核心输出	与输入B-rep模型在规定公差内，保持几何一致，且可编辑的参数化CAD模型，具有合理的特征树、草图与基准、尺寸参数和特征依赖关系；必要时说明所采用的建模方案及其中存在的不确定性。
评价标准	基础结果主要评价参数化模型能否成功生成，以及最终几何与输入B-rep的一致程度；进一步评价参数化结构是否合理、模型是否能够稳定编辑。

7.2 测试用例设计

7.2.1 测试样本选取

B-rep模型参数化重建任务的测试样本覆盖支架、连杆、泵体、叉轴、箱盖、发动机缸体，以及带轮支承、三爪卡盘和减速箱装配等结构。

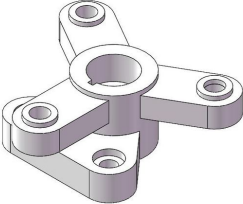
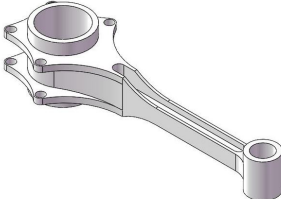
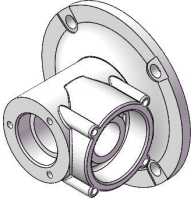
样本复杂度主要依据最终几何与结构关系划分，综合考虑曲面与孔槽的复杂程度、实体和零件数量、局部特征交互，以及装配组织难度，具体如表7-2所示。

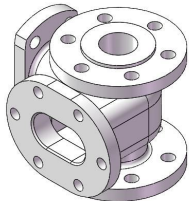
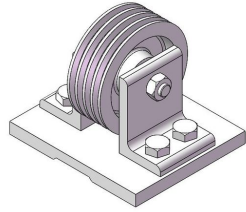
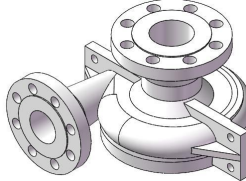
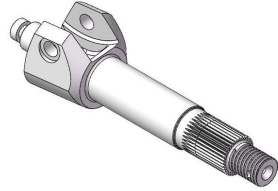
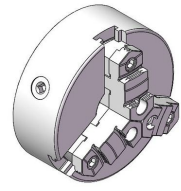
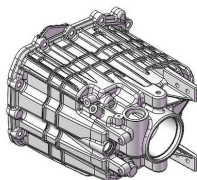
表7-2 B-rep模型参数化重建测试样本难度划分

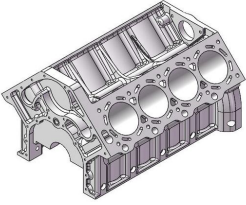
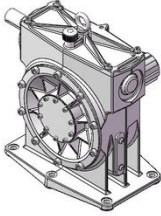
样本复杂度	选取标准	典型表现
简单	主体建模逻辑较直接，基础特征的作用范围清晰；即使包含较多重复孔或切除，特征之间仍以局部、相对独立的关系为主。	建模步骤间的依赖较少；基础拉伸/旋转/切除；局部孔特征、阵列、镜像。
中等	包含多种特征类型、多个草图或附加基准，存在一定的顺序选择、局部特征交互或多种合理重建方案；包括装配体。	多草图/多基准；圆角/倒角；阵列/孔特征；部分步骤依赖先前的草图或特征。
复杂	建模步骤较多且前后关联紧密，参考几何和特征交互较多，后续操作会覆盖或破坏部分早期特征边界；包括装配体。	大量旋转切除、孔特征或修饰特征；多项后续步骤依赖较早建立的特征；强交互。

基于上述难度分级，选取简单级别5个、中等级别3个和复杂级别3个样本，具体信息见表7-3。

表7-3 B-rep模型参数化重建测试样本清单

样本ID	CAD模型渲染图	模型几何面数	主要结构类型	复杂度	复杂度归类说明
B2P-01		60	支架、孔、阶梯截面	简单	规则多层结构，局部轮廓和轴向厚度关系较清楚。
B2P-02		83	连杆、双端孔、分层轮廓	简单	双端孔和杆身结构规则，主要尺寸关系清楚。
B2P-03		178	双向筒体、底法兰、沉孔、安装孔	简单	以规则筒体和局部孔系为主，轴向尺寸和孔特征的作用范围较清楚。

样本ID	CAD模型渲染图	模型几何面数	主要结构类型	复杂度	复杂度归类说明
B2P-04		153	多向流道、法兰、定位凸台、孔阵列	简单	以规则流道、法兰和重复安装孔为主，局部尺寸及其关联特征可分别检验。
B2P-05		348	带轮、轴、左右支承、紧固件	简单	16个实体，组成及位置关系明确，包含简单装配约束。
B2P-06		155	曲面泵腔、流道、法兰	中等	泵腔和法兰相交，曲面与内部通道增加重建难度。
B2P-07		380	叉头、轴段、槽孔	中等	叉头与轴段交接，局部槽孔和轮廓变化较多。
B2P-08		1628	卡盘体、卡爪、重复组件	中等	24个实体，重复实例较多，需组织零件并保存相对位置。
B2P-09		5639	筋板、孔腔、复杂曲面	复杂	5639个面，复杂曲面、筋板与孔腔交织，局部连通关系复杂。

样本ID	CAD模型渲染图	模型几何面数	主要结构类型	复杂度	复杂度归类说明
B2P-10		3422	缸孔、流道、密集孔系	复杂	缸孔与内部流道交错，径向轮廓和孔系细节密集。
B2P-11		10391	箱体、齿轮、轴、轴承组件	复杂	463个实体，零件实例密集，几何细节和装配层级复杂。

7.2.2 测试条件设计

(1) 主实验设计

主实验使用统一的Prompt对全部测试样本进行测试。每个样本以B-rep模型（STEP文件）作为重建输入，主实验仅改变测试样本，不改变Prompt，以保证不同样本和不同难度之间的结果可比较。具体Prompt内容为：“请根据提供的STEP模型，利用SolidWorks的建模API重建可编辑的参数化模型。要求：1.还原模型的几何形状，并建立合理的草图、特征和参数依赖关系。若为装配体，请保留独立零件及装配关系。2.如果重建的过程有进行多轮迭代，请按统一的评测方法，计算并记录每轮迭代的几何误差，输出误差变化曲线。无法计算的内容请如实说明。3.我会对每个模型进行约 10 次人工参数修改，检查参数是否生效、相关特征是否正确联动，以及模型是否出现异常。请提供主要参数的位置和作用。”

(2) 扩展实验设计

在主实验基础上，选取代表性复杂样本开展扩展实验，考察网络搜索权限的影响，关注GPT-6在缺少外部网络信息时，依据输入B-rep模型完成几何分析、建模历史推断与参数化重建的能力。扩展实验结果单独记录，不与主实验结果混合统计。

7.3 测试结果展示

本节分别展示主实验与扩展实验的测试结果。主实验重点评价GPT-6在统一Prompt下完成参数化重建的情况，并分析不同样本复杂度下的模型合法性、与参考CAD模型的几何一致性及可编辑性；扩展实验选取代表性复杂样本，通过关闭网络搜索，进一步考察模型对外部信息的依赖程度。两类实验结果分别展示与评价。

7.3.1 主实验结果

主实验中，GPT-6利用SolidWorks API进行参数化重建，生成包含草图、尺寸参数和建模操作的工程文件。本节先汇总11个样本的测试结果，再展示参考CAD模型与重建模型的同视角对比，结合几何测量和参数修改分析各用例的表现。

(1) 总体概览

11个样本均采用第7.2.2节规定的统一Prompt进行重建。几何一致性以双向表面均方根距离RMS为指标。先把参考和重建模型转化成三角网格，分别在参考与重建网格上按三角面面积均匀采样，计算各点到对方网格的最近距离，越小表示整体表面越接近。

先对重建模型进行实体有效性检查，未通过实体有效性检查的模型记为“无法编辑”，不进行参数修改测试。对通过实体有效性检查的模型，在SolidWorks中每个进行10次单参数修改。单参数修改是指每次只主动改变一个尺寸参数，其余输入参数保持不变，相关特征按已有依赖关系联动，每次测试后恢复原值。只有参数修改生效、模型成功重算、几何产生预期变化、相关结构正确联动，未出现孔口封闭、结构破损或组件错位等问题，且能恢复原状态，才计为通过，具体结果见表7-4。

表7-4 B-rep模型参数化重建总体结果

样本ID	难度	实体有效性	几何一致性 RMS / mm	可编辑性
B2P-01	简单	1/1通过	0.000698	10/10通过
B2P-02	简单	1/1通过	0.000158	10/10通过
B2P-03	简单	1/1通过	0.05565	4/10通过
B2P-04	简单	1/1通过	0.04671	8/10通过
B2P-05	简单	16/16通过	0.00175	10/10通过
B2P-06	中等	2/3通过	2.9687	无法编辑
B2P-07	中等	1/1通过	0.5304	10/10通过
B2P-08	中等	23/24通过	0.4193	无法编辑
B2P-09	复杂	20/23通过	9.9593	无法编辑
B2P-10	复杂	0/1通过	2.6181	无法编辑
B2P-11	复杂	262/278通过	2.2087	无法编辑

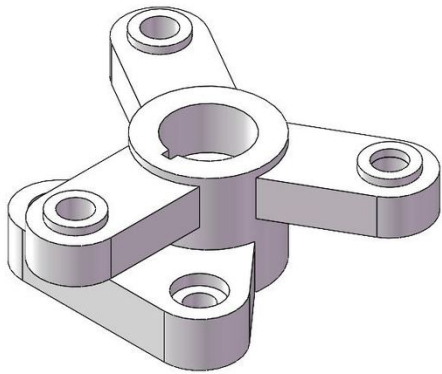
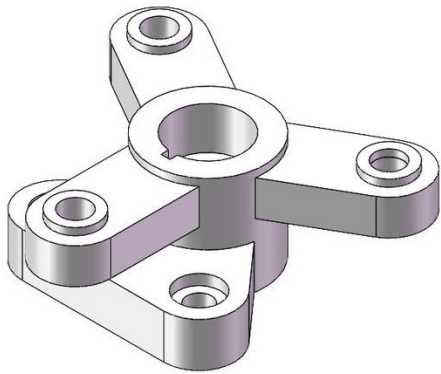
表7-5为GPT-6重建各样本具体所用时间。

表7-5 各样本重建用时

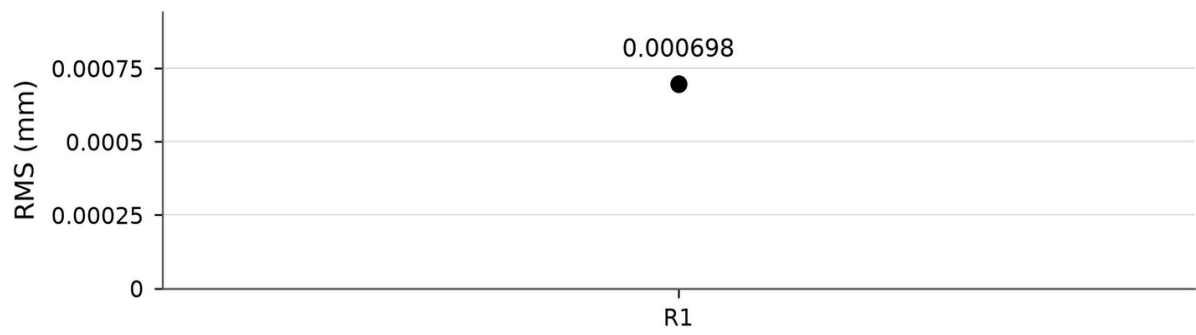
样本ID	B2P-01	B2P-02	B2P-03	B2P-04	B2P-05	B2P-06	B2P-07	B2P-08	B2P-09	B2P-10	B2P-11
用时	1 分钟	1 分钟	9 分钟	16 分钟	11 分钟	13 分钟	6 分钟	21 分钟	33 分钟	5 分钟	1 时 52 分钟

(2) 用例分析

B2P-01（简单）：三臂法兰轴

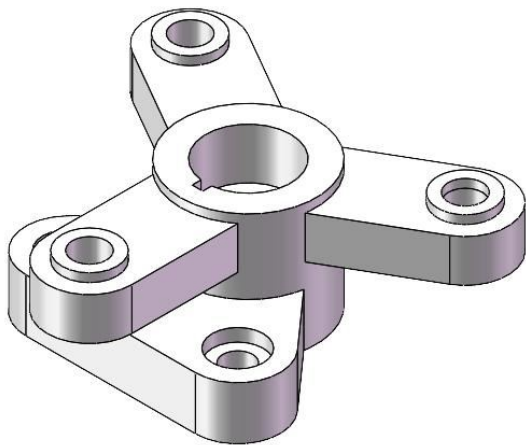
参考CAD模型预览	GPT-6重建模型预览
 面数：60	 面数：61

B2P-01 几何误差曲线

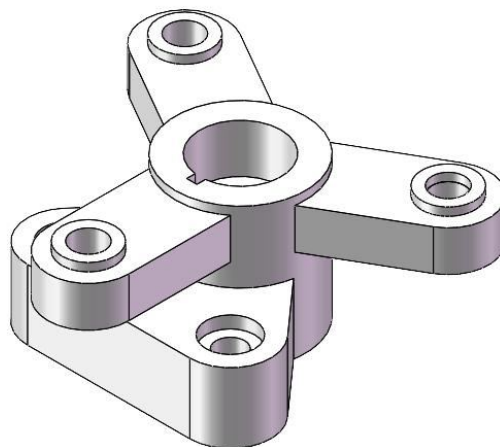


代表性参数修改：BasePlate: 19 → 28.5 mm; 通过

修改前



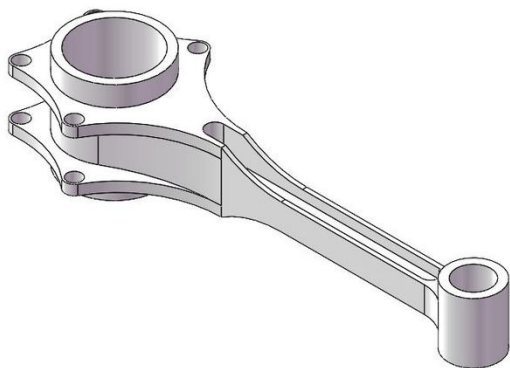
修改后



结果分析：生成的模型实体有效，GPT-6生成过程中仅进行了一次迭代，恢复了三臂法兰轴的主要结构，整体几何与参考模型高度接近。10项参数修改测试全部通过，修改后的几何能够更新并恢复。

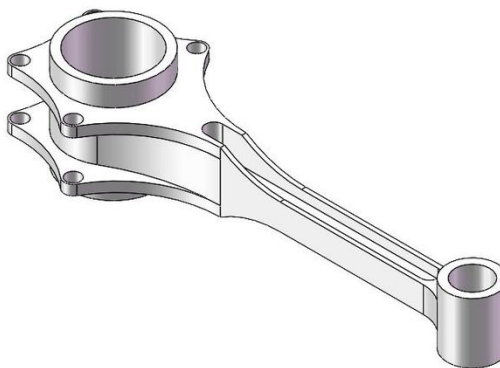
B2P-02（简单）：连杆

参考CAD模型预览



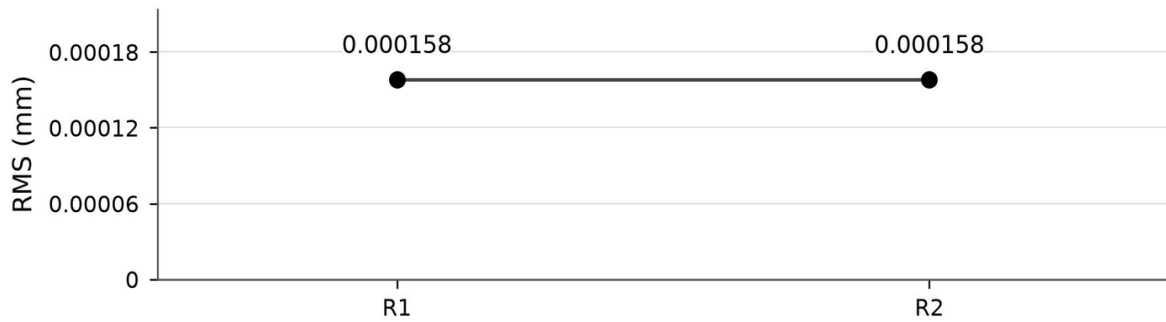
面数：83

GPT-6重建模型预览



面数：83

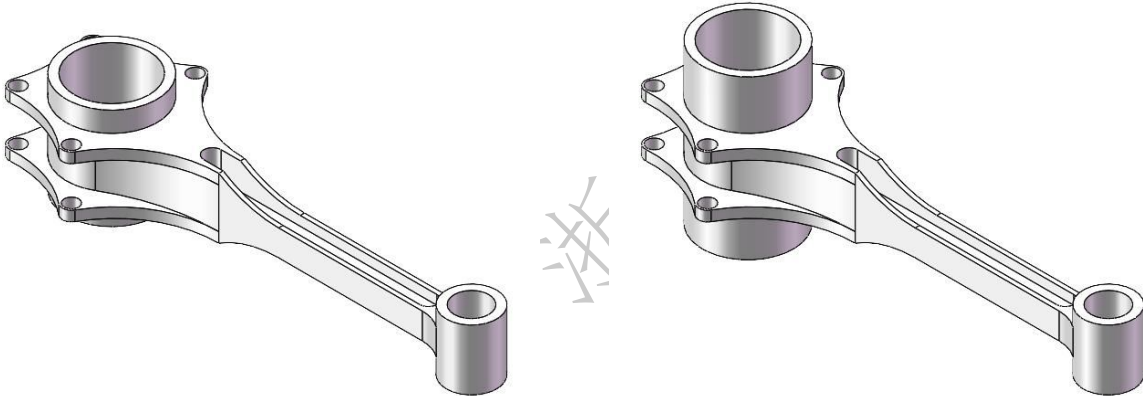
B2P-02 几何误差曲线



代表性参数修改: BigEndWidth: 70 → 109.9 mm; 通过

修改前

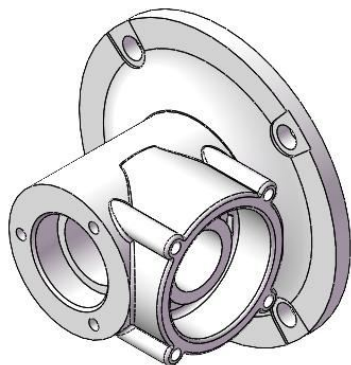
修改后



结果分析: 生成的模型实体有效, 两轮迭代的RMS均为0.000158 mm, 保存版本间未见明显几何改善, 连杆主体与两端孔的整体几何接近参考模型。10项参数修改均通过, 修改后的几何能够更新并恢复。

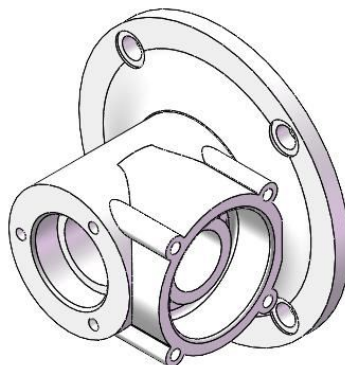
B2P-03（简单）：法兰壳体

参考CAD模型预览

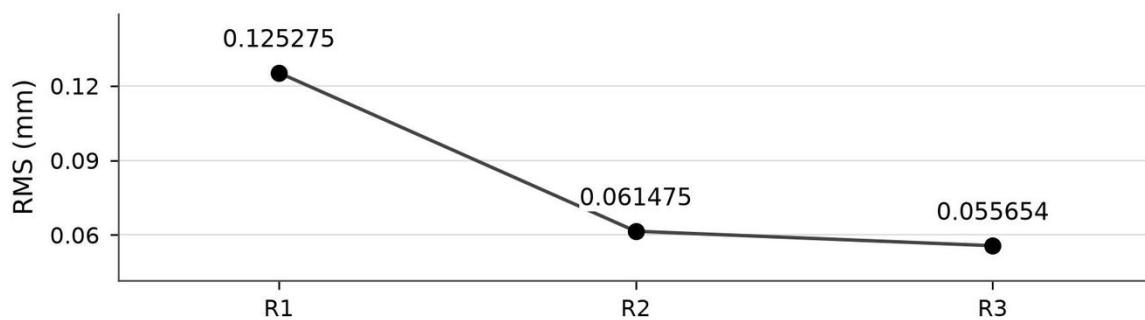


面数: 178

GPT-6重建模型预览

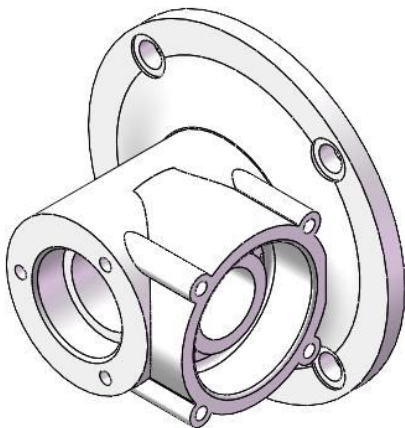


面数: 114

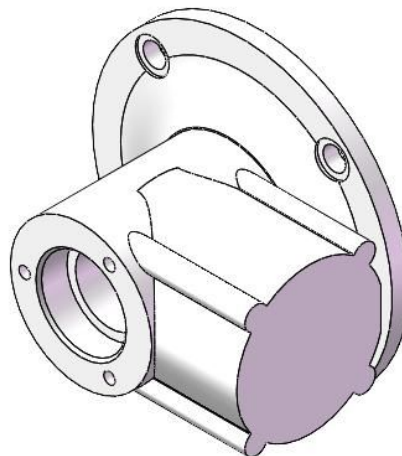
B2P-03 几何误差曲线

代表性参数修改: SideHousing_Boss深度: 36 → 54 mm; 失败

修改前



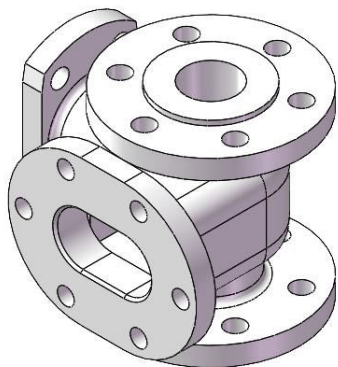
修改后



结果分析：生成的模型实体有效，三轮迭代中RMS由0.1252 mm降至0.0556 mm。模型较接近地恢复了双向筒体、底法兰与主要孔系。10项参数修改中，4项通过、6项失败，图示将侧壳深度由36增至54 mm，原孔口被封住。其余失败涉及侧部结构截断、安装孔消失、沉孔台阶消失或相关特征重建错误。

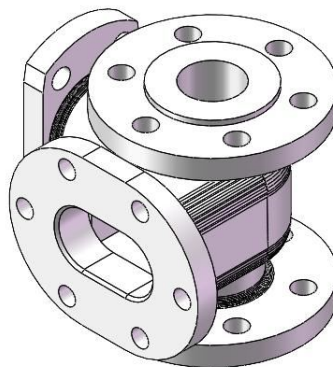
B2P-04（简单）：多通道法兰壳体

参考CAD模型预览



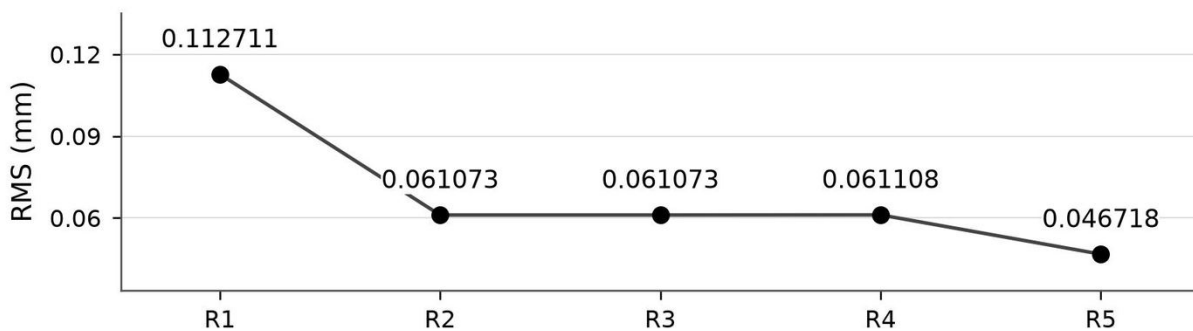
面数: 153

GPT-6重建模型预览

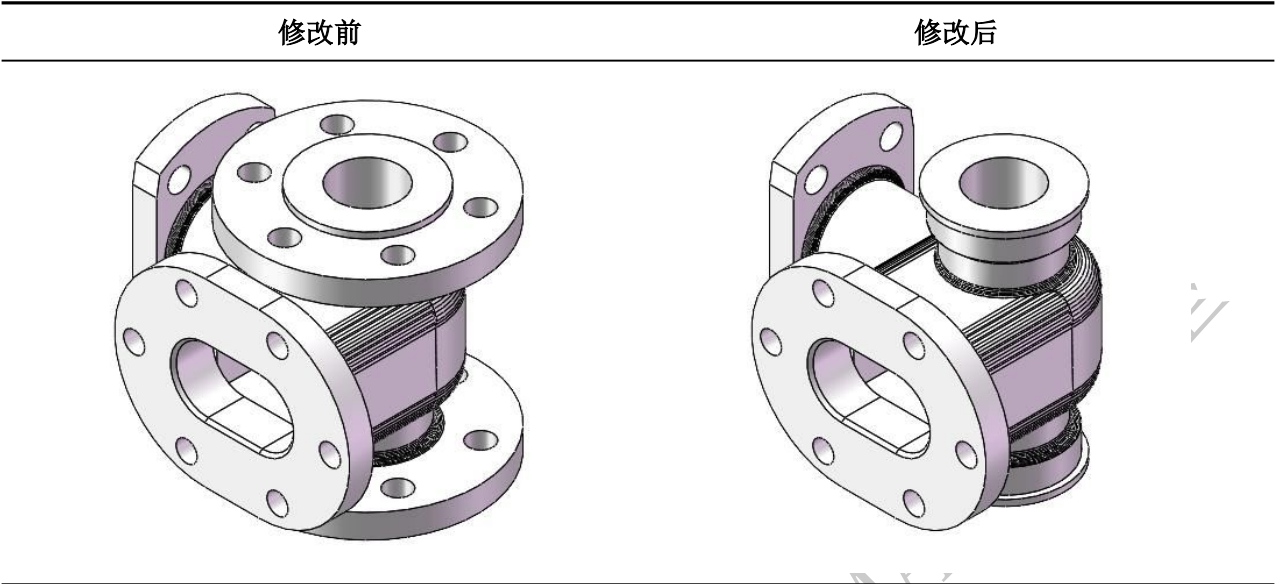


面数: 1401

B2P-04 几何误差曲线

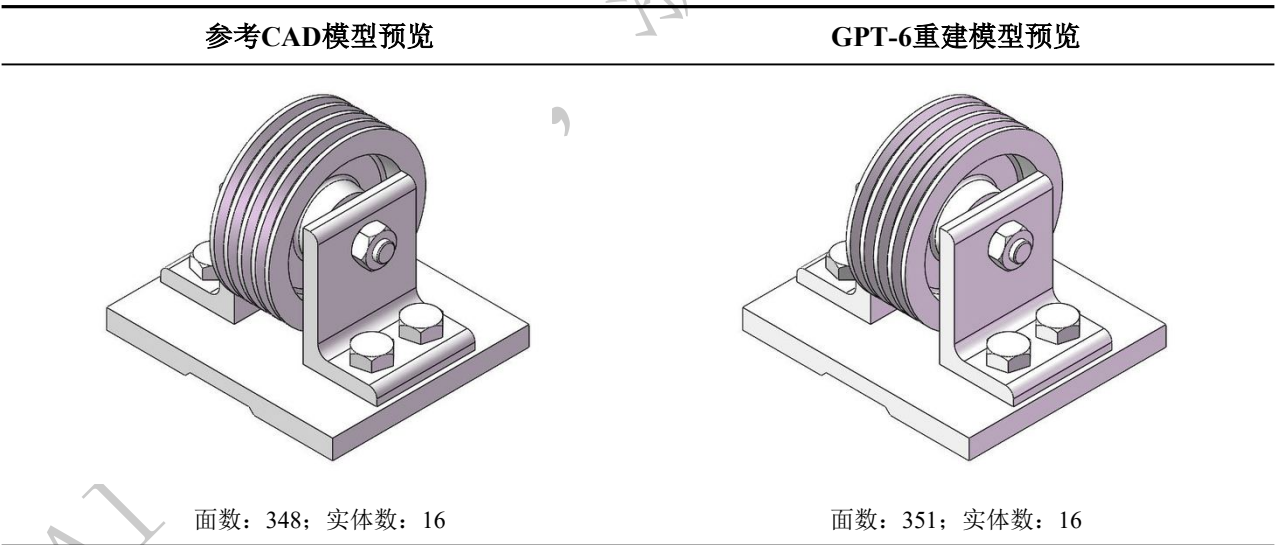


代表性参数修改: MainFlangeOD: 146 → 73 mm; 失败

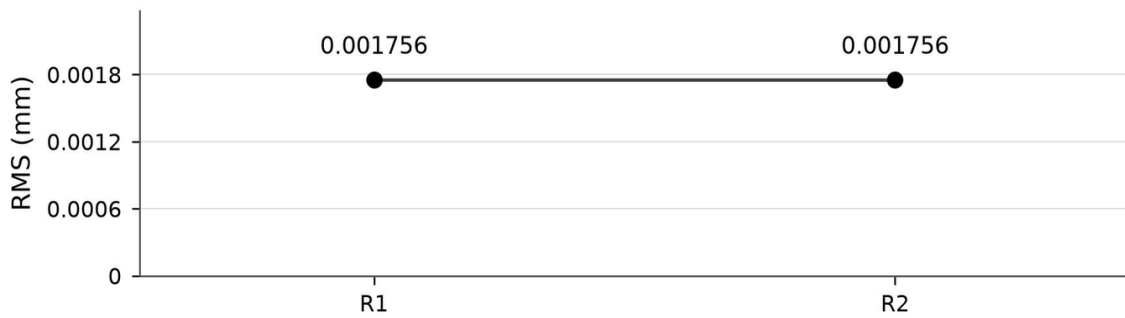


结果分析: 生成的模型实体有效, 五轮迭代中RMS由0.1127 mm降至0.04671 mm。10项参数修改中, 8项通过、2项失败。图示将主法兰外径由146减至73 mm, 安装孔位置未随外径调整, 孔系超出缩小后的法兰并消失。

B2P-05 (简单): V型皮带轮和轴承支撑组件

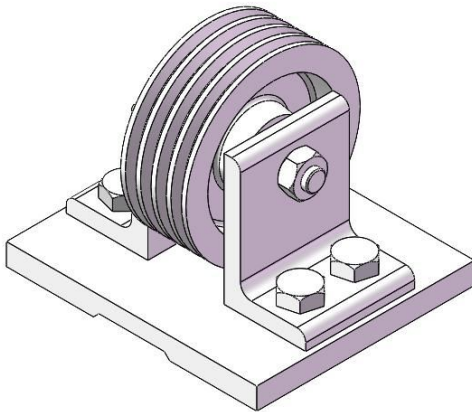


B2P-05 几何误差曲线

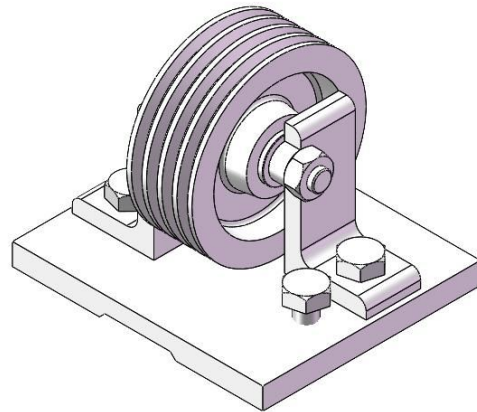


代表性参数修改: 15_Bearing_Right / BracketWidth: 100 → 50 mm; 通过

修改前



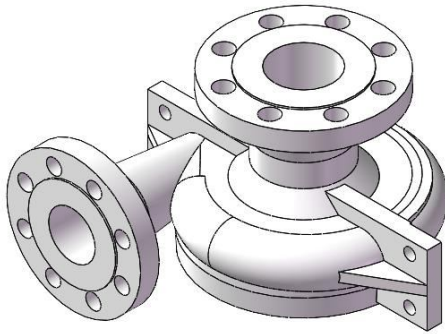
修改后



结果分析: 生成的装配体中16个零件均实体有效, 两轮迭代的RMS均为0.001756 mm, 参考与重建整体STEP均为16个实体, 主体结构在测量范围内较接近。10项参数修改均通过, 修改后的几何能够更新并恢复。图示将右支承宽度由100减为50 mm, 其他组件结构与状态均保持正常。

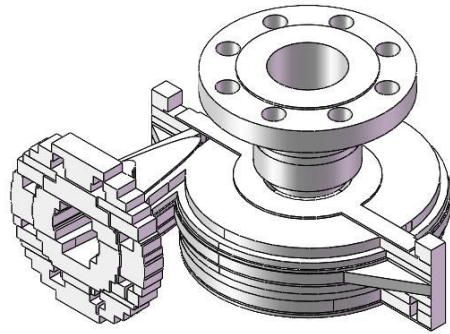
B2P-06（中等）：离心泵壳体

参考CAD模型预览

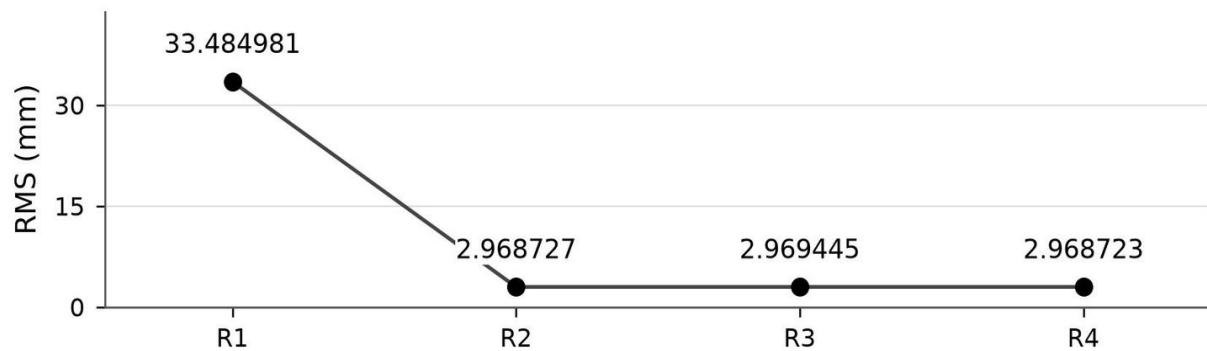


面数: 155

GPT-6重建模型预览



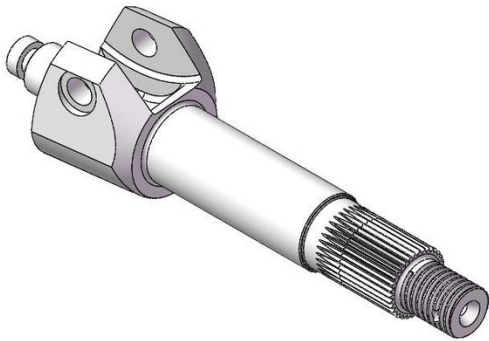
面数: 1081

B2P-06 几何误差曲线

结果分析：参考模型仅有1个实体，重建模型中有3个实体，重建模型中2个实体有效，1个实体无效。四轮迭代中，RMS由33.485 mm降至约2.9687 mm后趋于平稳，模型的泵腔、流道及连通结构仍存在偏差。由于存在未通过有效性检查的实体，该模型记为“无法编辑”，不进入参数修改测试。

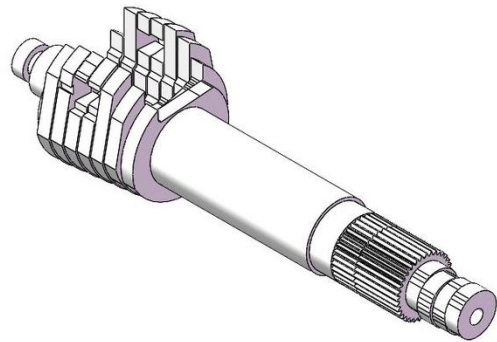
B2P-07（中等）：控制杆轴

参考CAD模型预览

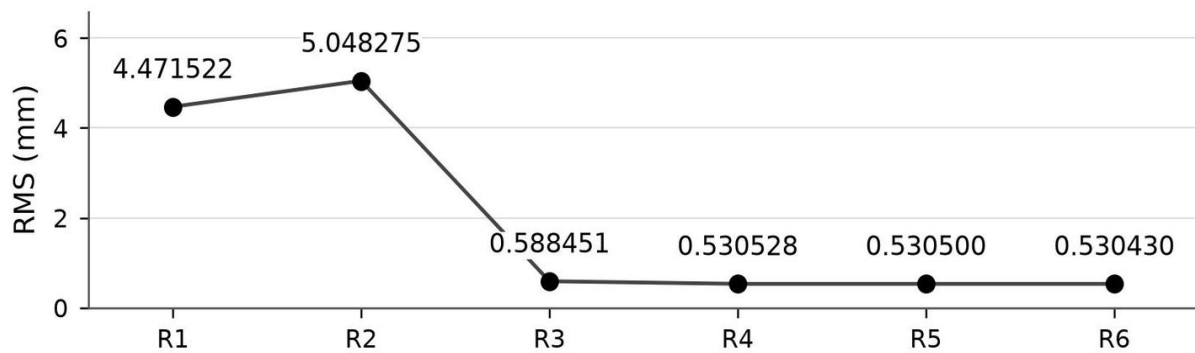


面数: 380

GPT-6重建模型预览

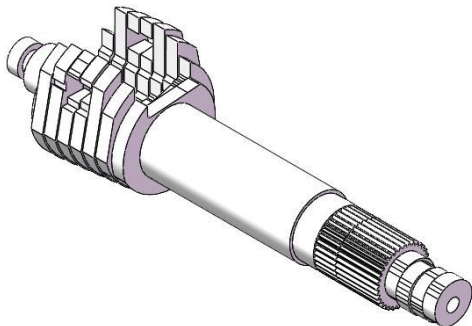


面数: 825

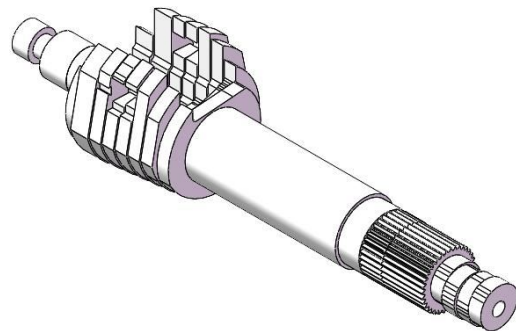
B2P-07 几何误差曲线

代表性参数修改: Zone1Length: 44.75 → 67.125 mm; 通过

修改前



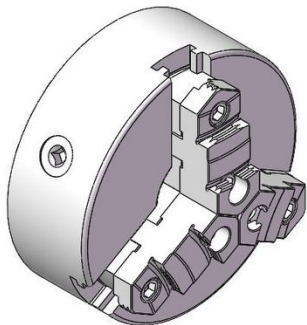
修改后



结果分析：生成的模型实体有效，六轮迭代中，RMS先由4.4715 mm升至5.0483 mm，随后降至约0.5304 mm，模型的叉头与轴段交接等局部结构仍有偏差，仅是与参考模型在视觉上相似。10项参数修改均通过，修改后的几何能够更新并恢复。

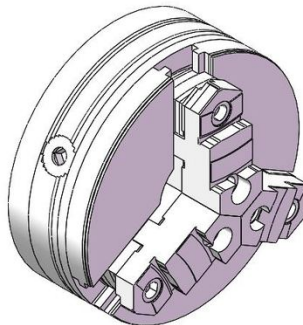
B2P-08（中等）：车床卡盘

参考CAD模型预览



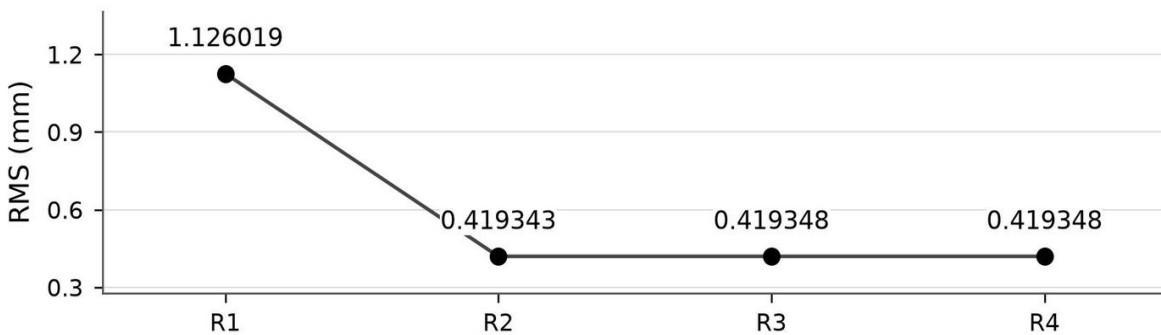
面数：1628；实体数：24

GPT-6重建模型预览



面数：6544；实体数：24

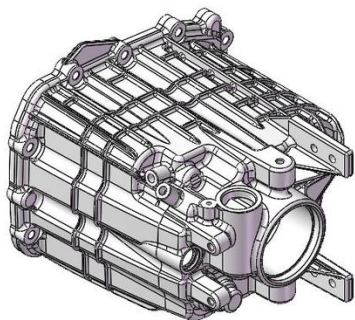
B2P-08 几何误差曲线



结果分析：重建模型中23个实体有效，1个实体无效，文件中还含有371个未归属Solid的游离面。四轮迭代中，RMS由1.1260 mm降至约0.4193 mm后基本保持稳定。由于存在未通过有效性检查的实体，该模型记为“无法编辑”，不进入参数修改测试。

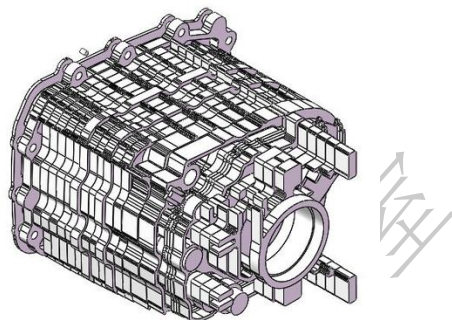
B2P-09（复杂）：变速箱壳盖

参考CAD模型预览

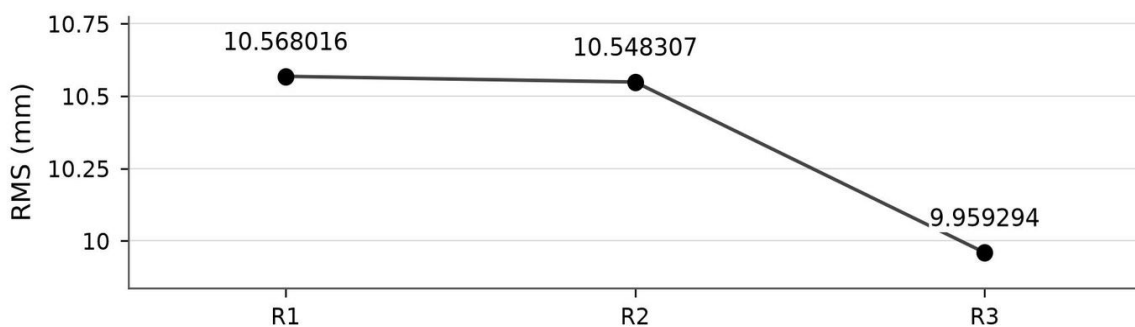


面数：5639

GPT-6重建模型预览



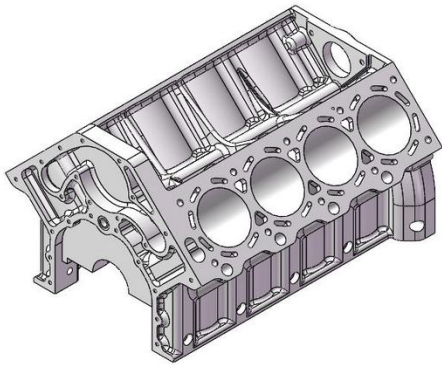
面数：8466

B2P-09 几何误差曲线

结果分析：参考模型中仅有3个实体，重建模型生成了23个实体，其中20个有效，3个无效。三轮迭代中，RMS由10.5680 mm降至9.9593 mm，整体偏差仍大。重建采用多段拉伸形成的简单面近似箱盖的复杂曲面，外表面出现密集分段。由于存在未通过有效性检查的实体，该模型记为“无法编辑”，不进入参数修改测试。

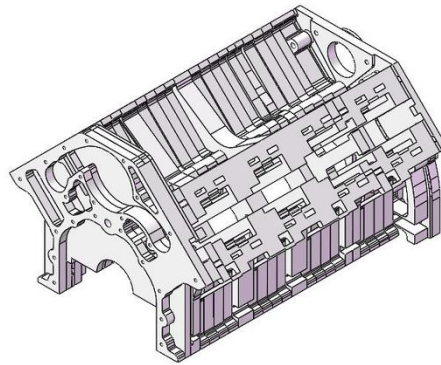
B2P-10（复杂）：V8型发动机缸体

参考CAD模型预览

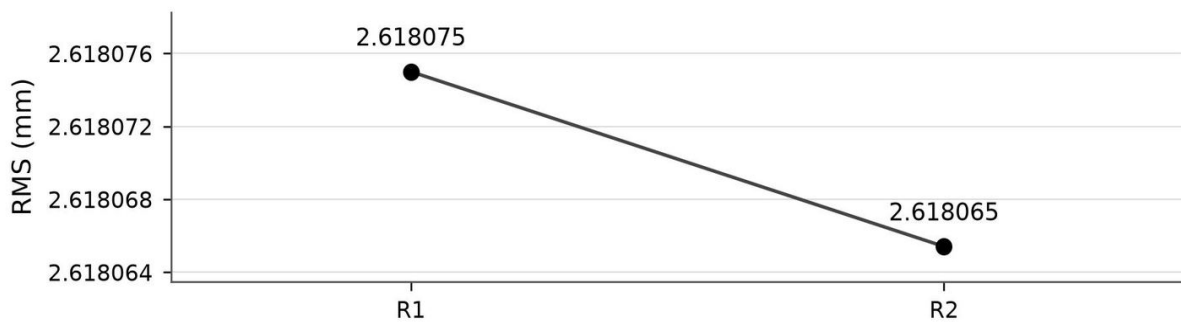


面数: 3422

GPT-6重建模型预览



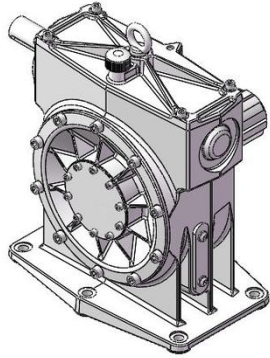
面数: 4830

B2P-10 几何误差曲线

结果分析：生成的模型实体无效。两轮迭代的RMS均约为2.6181 mm，优化较小，重建模型与参考模型的整体结构差异很大，重建采用分截面拉伸，用多段简单面堆叠近似缸体曲面和孔道轮廓，面数由3422增至4830，缸孔与流道未完整恢复。由于未通过实体有效性检查，该模型记为“无法编辑”，不进入参数修改测试。

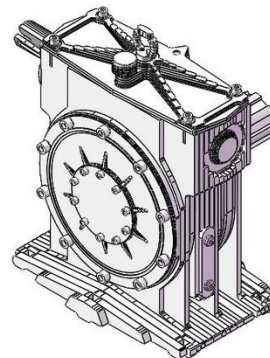
B2P-11（复杂）：蜗轮行星齿轮箱

参考CAD模型预览



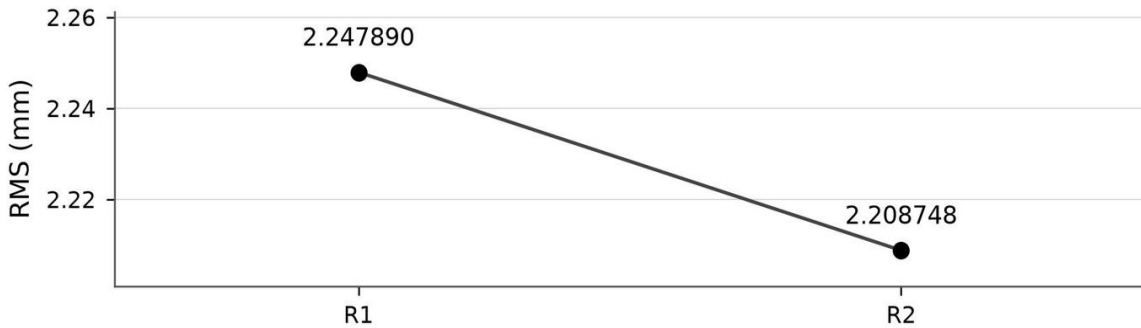
面数：10391；实体数：463

GPT-6重建模型预览



面数：55640；实体数：278

B2P-11 几何误差曲线



结果分析：参考模型有463个实体，重建模型中仅有278个实体，其中262个实体有效、16个实体无效。两轮迭代中RMS由2.2479 mm降至2.2087 mm，改善较小。重建零件采用分截面拉伸，以多段简单面近似壳体及其他部件的曲面，面数由10391增至55640。由于存在未通过有效性检查的实体，该模型记为“无法编辑”，不进入参数修改测试。

7.3.2 扩展实验结果

扩展实验对7.2.2节选取的代表性复杂样本关闭网络搜索，并与对应的主实验结果进行对照。除网络搜索权限外，两种条件使用相同的输入模型、Prompt、模型配置和输出要求，以观察关闭搜索后重建流程及结果是否发生明显变化。结果见表7-6。

表7-6 主实验与无网络条件的结果对比

样本ID	测试条件	实体有效性	RMS / mm	可编辑性
B2P-09	主实验	20/23通过	9.9593	无法编辑
B2P-09	无网络	1/3通过	0.6074	无法编辑
B2P-10	主实验	0/1通过	2.6181	无法编辑
B2P-10	无网络	0/2通过	1.6184	无法编辑
B2P-11	主实验	262/278通过	2.2087	无法编辑
B2P-11	无网络	262/278通过	2.2067	无法编辑

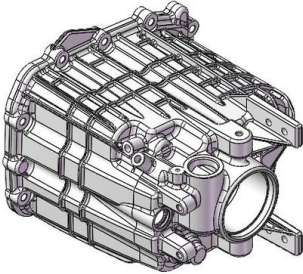
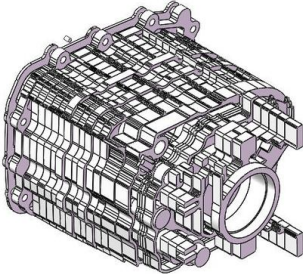
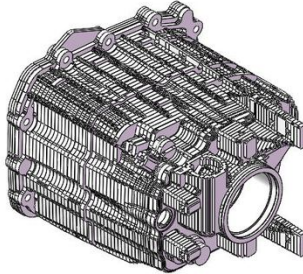
表7-7为GPT-6重建各样本具体所用时间。

表7-7 各样本重建用时

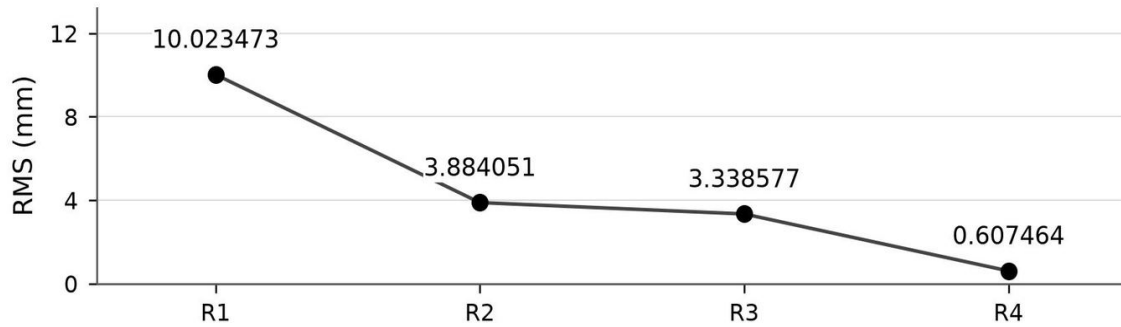
样本ID	B2P-09主实验	B2P-09无网络	B2P-10主实验	B2P-10无网络	B2P-11主实验	B2P-11无网络
用时	33分钟	1时0分钟	5分钟	48分钟	1时52分钟	1时15分钟

扩展实验具体用例分析如下。

B2P-09（变速箱壳盖）（无网络条件）

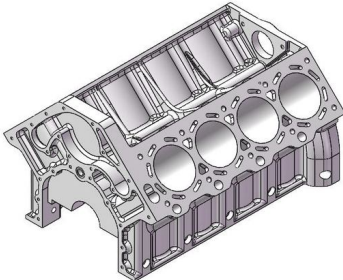
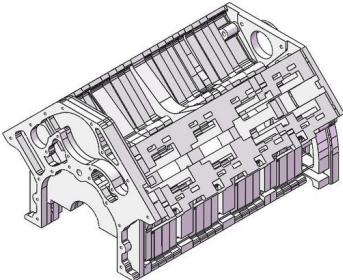
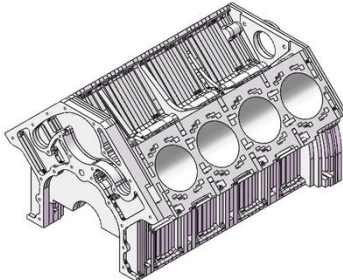
参考CAD模型	主实验重建模型	无网络重建模型
		
面数：5639；实体数：3	面数：8466；实体数：23	面数：21567；实体数：3

B2P-09 无网络条件几何误差曲线

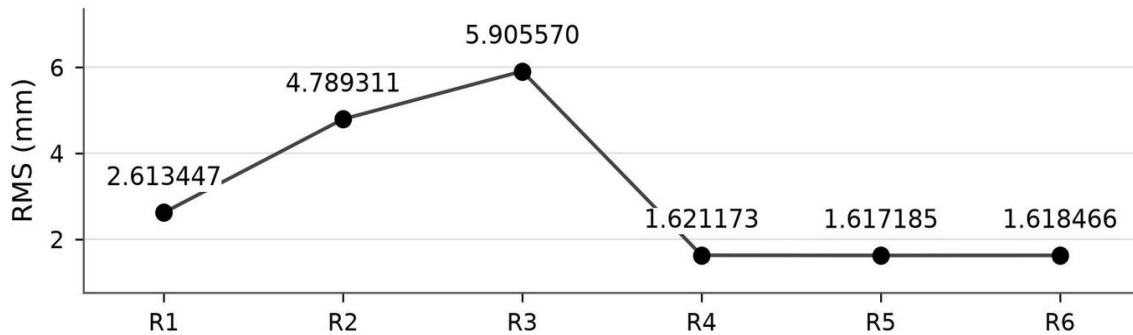


结果分析：实体数量与参考模型一致，实体有效性检查中，1个实体通过、2个未通过，因此记为“无法编辑”，不进入参数修改测试。四轮迭代中RMS从10.0234 mm降至0.6074 mm。无网络重建的复杂曲面仍由多段简单面逐层近似，面数增至21567，分段更加密集。几何误差较主实验减小，但实体有效性仍未全部通过。

B2P-10 (V8型发动机缸体) (无网络条件)

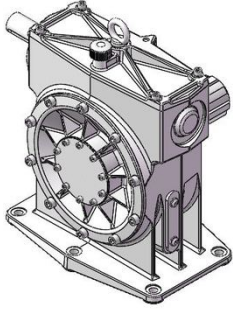
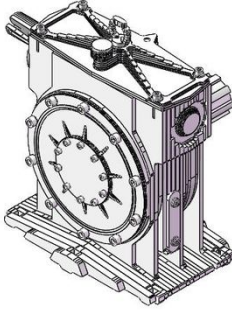
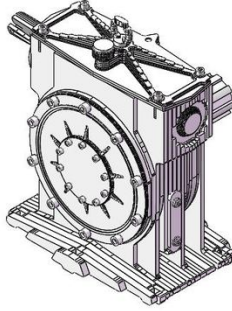
参考CAD模型	主实验重建模型	无网络重建模型
		
面数：3422；实体数：1	面数：4830；实体数：1	面数：11746；实体数：2

B2P-10 无网络条件几何误差曲线

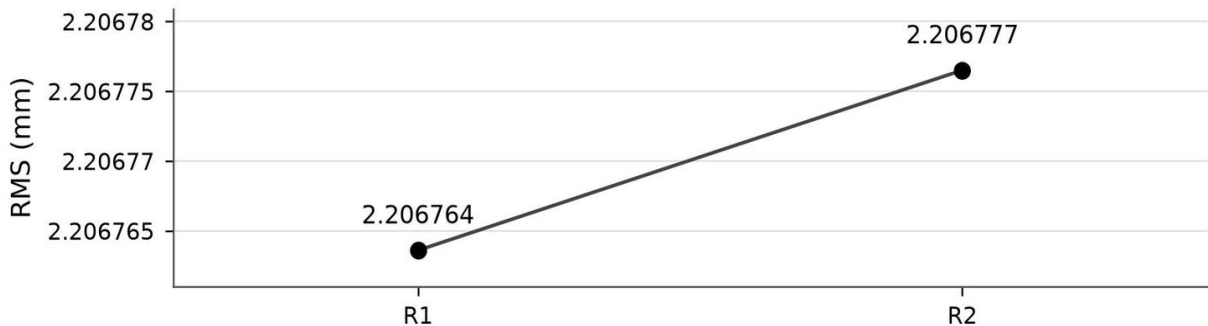


结果分析：重建模型的2个实体均未通过有效性检查，因此记为“无法编辑”，不进入参数修改测试。六轮迭代中，RMS先从2.6134mm升高至5.9055mm，随后降至1.6184 mm，缸体曲面仍由多段简单面堆叠近似。

B2P-11（蜗轮行星齿轮箱）（无网络条件）

参考CAD模型	主实验重建模型	无网络重建模型
		
面数：10391；实体数：463	面数：55640；实体数：278	面数：55637；实体数：278

B2P-11 无网络条件几何误差曲线



结果分析：实体有效性检测中，262个实体通过、16个实体未通过，因此记为“无法编辑”，不进入参数修改测试。两轮迭代中RMS为2.2067和2.2067mm，基本不变，重建仍采用多段简单面近似曲面，整体几何误差与主实验接近。

综合判断：无网络条件下，GPT-6仍能够生成原生参数化模型。与主实验相比，模型质量未呈现一致且明显的优劣差异，实体有效性、几何一致性和参数可修改性仍存在明显局限。在本次样本与测试条件下，尚未观察到网络搜索对参数化重建能力的明确影响。

7.4 GPT-6推理过程分析

基于主实验中对所有样本的测试，以及多次分析hack，GPT-6对B-rep模型参数化重建任务可能采用如下推理流程：

① 读取几何信息并制定重建方案

分析脚本提取几何参数，检查轮廓、孔槽及实体相交关系



② 建立参数、草图与建模特征

建模脚本将尺寸参数与草图、特征关联，生成可编辑模型



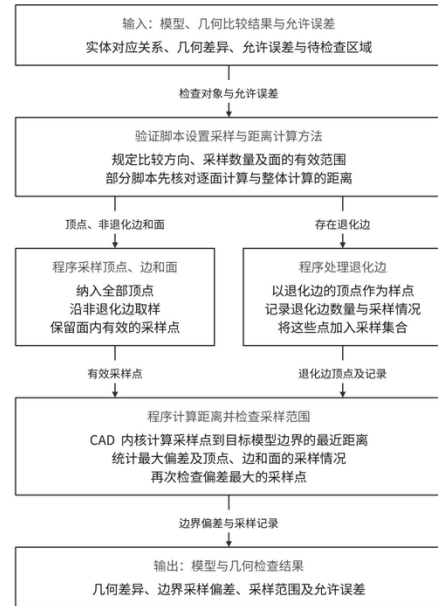
③ 比较模型几何并检查误差

程序比较输入模型与重建结果，按允许误差检查差异



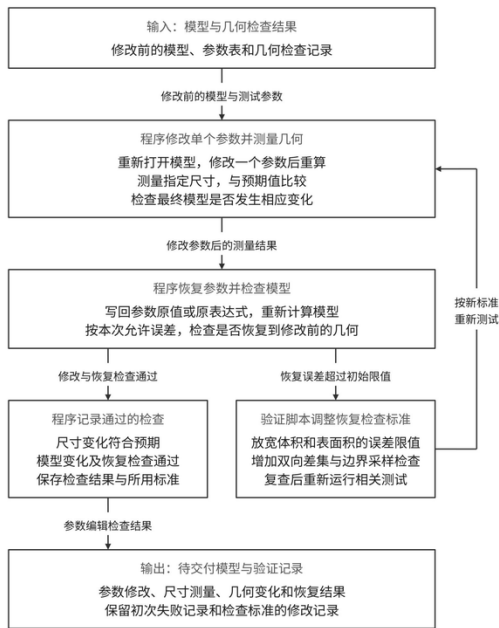
④ 检查边界偏差与采样范围

程序在顶点、边和面上采样，检查重建模型的边界偏差



⑤ 验证参数编辑与模型恢复

程序逐项修改参数，检查几何变化及恢复结果



⑥ 保存模型并核对文件版本

程序保存显示设置，确认交付文件与验证记录对应



7.5 能力边界

在本章的B-rep模型参数化重建测试中，GPT-6能够读取输入模型的几何与拓扑信息，建立草图、尺寸参数及拉伸、旋转、切除等特征，并生成可编辑的零件和装配体。三臂法兰轴、连杆及带轮支承组件的主要形状得到恢复；对于泵体、叉轴和复杂壳体，GPT-6也能通过多轮建模调整减小整体几何误差。然而，GPT-6对复杂曲面的重建和参数依赖关系的建立仍不完整。部分模型将连续曲面拆为多个截面，通过堆叠简单面来近似原形状。变速箱壳盖的外表面因此出现密集分段，发动机缸体的孔道轮廓也未完整恢复。参数修改也暴露出关联特征关系的缺失，11个样本中，只有6个模型通过实体有效性检查，具备编辑测试条件，其中5个简单、1个中等。对这6个模型各进行10次单参数修改，共52次通过，10次修改全部通过的模型只有4个，占全部样本的36.4%，其中3个简单、1个中等。B2P-03法兰壳体加深后孔口被封住，B2P-04法兰缩小后安装孔消失，说明相关特征未能随主体尺寸正确更新。

GPT-6已具备根据B-rep模型建立原生参数模型并编辑部分尺寸的能力，但尚不能稳定保证复杂曲面和装配关系同时正确。部分重建实体未通过有效性检测，卡盘中还存在游离面。无网络组仍采用简单面分层近似复杂曲面的方式，现有三个对照样本不足以确认网络搜索对重建质量的影响。本次测试主要覆盖单参数修改，尚未验证多参数组合、特征增删和完整装配功能。

8 综合分析

前述各章分别考察了 GPT-6 在零件分类、特征识别、模型生成/重建等六类任务中的表现。本章结合不同任务的结果，进一步讨论当前以 GPT-6 为代表的 AI 模型已经能够承担哪些 CAD 任务，哪些要求仍难以满足，以及智能体式（Harness Engineering 和 Skills）的迭代优化带来了什么能力。

零件分类与特征识别的结果表明，GPT-6 已经能够结合具体几何信息和机械设计知识，对 CAD 模型作出语义层面的判断。例如，GPT-6 能够根据偏心轴段及曲柄的连接识别四拐曲轴，根据线圈与力臂的布置情况识别双扭转弹簧（PC-11、PC-12）。在轴承案例中，GPT-6 确认了滚子轴承大类，同时注意到滚子主体为圆柱面，并据此对其是否属于标准圆锥滚子轴承给出了意见（PC-09）。GPT-6 还能够进行局部几何的划分及特征之间关系的识别：FR-04 正确识别了多级嵌套腔槽、相交通道，以及被通道截断的安装孔。这些结果说明，GPT-6 已经能结合几何信息、机械设计知识及网络检索内容，辨别零件类别、识别特征，并给出相应的判断依据。

不过，上述测试结果也表明，即使局部几何识别正确，仍可能在判断整体用途时出错。在取消候选类别提示后，GPT-6 虽然在 PC-10 中识别出了弹簧和阀杆，却据此将管路接头判断为止回阀，将内部阀机构的功能误认为整个接头的用途。特征划分也存在类似错误：FR-06 将空间上分离的两段孔壁合并为一个孔，并将实体外侧的圆柱面误识别为孔。可见，系统已经能够利用几何信息进行分析，但哪些面应归为同一特征、局部特征在整体中承担什么作用，仍容易出现错误判断。

在模型生成方面，成功案例包括具有复杂内部孔道的实体和包含多组机构的装配体。在给定详细几何特征描述的条件下，TC-05 生成了单一有效的 V8 缸体实体，水套保持连通，并与缸孔隔离；TC-06 完成了卧式单缸蒸汽机主传动机构与配气机构的建模和装配，活塞杆、十字头、曲柄连杆及偏心配气机构的组成和相对位置符合要求。这些结果表明，在给定几何要求下，GPT-6 已经能够完成复杂内部几何的构造，以及多个零件的定位与装配。

功能需求建模测试进一步展示了 GPT-6 从需求出发生成设计方案的能力。在 TRC-03 中，GPT-6 围绕指定叶轮设计泵壳，构造了连通流道、管路接口和后部拆卸开口，为紧固工具及转子拆卸预留了空间，并完成了给定工况下的水头损失与承压计算。在 TRC-04 的气动双指夹爪设计中，GPT-6 采用气缸驱动、齿轮齿条同步的平行夹持方案，通过导向轴约束夹指运动。该方案不同于参考模型的连杆机构，但仍满足测试列出的功能要求。这些工作涉及机构选型、接口设计、空间布置和分析计算，说明 GPT-6 已经能够在部分任务中作出具体的设计选择，并将其落实为 CAD 模型。

测试结果也表明，复杂曲面、薄壁和局部过渡的准确生成与重建，在不同任务中仍反复出现困难。工程图重建中的 D2C-A03 未能准确恢复复杂曲面与薄壁区域；图像重建中的 I2C-03 采用多截面直纹放样近似控制臂侧壁，未完整恢复光顺曲面及根部圆角；B-rep 模型参数化重建中的 B2P-09 和 B2P-10，则采用分截面拉伸近似壳体曲面及孔道轮廓，存在较大几何偏差或无效实体。这些案例的主要外形和部分孔槽已经能够生成，困难集中在复杂曲面、连接过渡等复杂几何细节。

即使主要外形已经生成，局部孔道、实体构成和装配关系也可能不符合要求。D2C-P05 生成了有效的缸体实体，却将部分通孔建成盲孔。I2C-03 的三视图测试生成了控制臂主体、衬套和耳板，但三者是局部重叠的独立实体，是无效模型。TC-04 的夹爪包含电机、齿轮和相关机构零件，电机轴与小齿轮中心孔之间却留有径向间隙，也未设置紧定螺钉等固定结构，无法可靠传递转矩。功能需求建模中的直升机也出现了类似问题：TRC-07 变距拉杆未连接到倾斜盘，云台电机未固定到支板。这些缺陷涉及的几何误差很小，却直接影响模型功能。也就是说，GPT-6 能够完成相当复杂的几何和部件布置，却经常遗漏决定功能的局部要求。

对于上述任务中的失败案例，不能全部归因于输入信息不足。图像缺少尺度、三视图未完整显示内部结构，确实会限制几何恢复，但 B-rep 模型参数化重建任务提供的 STEP 文件已经包含全部几何信息，重建结果仍出现明显近似（B2P-09 和 B2P-10 采用分截面拉伸近似壳体曲面及孔道轮廓，存在较大几何偏差或较多无效实体），近一半案例的参数化重建模型为错误模型。实际上，完整的几何信息需要经过提取和分析，再转化为合适的建模特征与操作，才能得到准确的重建结果。现有测试说明，即使目标几何完整，GPT-6 尚未能够充分利用这些信息完成复杂模型的准确重建。

参数化重建又将问题推进到了设计修改环节。B2P-03 的法兰壳体在初始状态下实体有效，双向表面距离的均方根值约为 0.0556 mm，但一旦进行参数修改，模型便失效，例如，将侧壳深度由 36 mm 增加至 54 mm 后，原有孔口被封闭；其他修改还出现了安装孔、沉孔台阶消失或相关特征重建错误。B-rep 描述了模型的形状，却不能唯一确定应采用怎样的特征历史，以及哪些尺寸和零件需要随修改联动。测试结果说明，GPT-6 基本不能保证后续参数修改符合预期。

理解上述能力与局限，还需要考察 GPT-6 完成 CAD 任务所采用的技术范式。对本报告中大部分案例，GPT-6 采用了智能体式工作流程：将模型生成、结果检查与反馈修正组织为多轮迭代。例如，TC-05 曾出现水套错误贯通中央凹腔、纵向孔道被封闭的问题，经过迭代修改后才实现所需的连通与隔离；TRC-02 也通过调整止推垫片、限位件和钳口紧固部位，消除了活动部件的局部干涉。这些修改说明 GPT-6 已经能够在迭代过程中对已有模型中的具体问题进行分析修正。

不过，某些问题得到修正，并不意味着迭代结束时要求就能满足。例如，TC-07 所列几何要求的满足率达到 100%，最终模型仍含有非流形连接和自交曲面，未通过整体实体有效性检查。B2P-08 的均方根值由 1.1260 mm 降至约 0.4193 mm，但后续 10 项参数修改测试基本失败。迭代优化的效果因而需要结合不同要求分别判断，某项指标的改善尚不足以说明任务已经完成。

此外，值得注意的一个现象是，迭代过程中误差曲线经常出现波动，表明修改并不总能带来预期的改善。这提示 AI 模型对误差来源、修改对象及相应操作的判断仍不够准确。同时，CAD 模型中的几何特征与参数相互依赖，一处修改可能同时影响多个设计要求，在修正原有问题时引入新的偏差。因此，当前迭代优化的困难在于准确确定需要修改的部位及操作范围，并判断修改对相关特征和设计要求的影響，从而有针对性地解决问题。

迭代修正还带来了较大的时间开销。测试中，多类建模与重建任务耗时数十分钟。其中，复杂文本建模样本的平均用时为 1 小时 37 分 30 秒，该类样本的最长用时为 3 小时 17 分钟。从执行过程看，部分任务需要反复构建、检查和修正模型，才能逐步满足要求。虽然这些任务已取得较好的建模结果，但较长的执行时间使其难以适应设计过程中快速反馈、连续修改和方案比较的需要。当前系统已展示出完成复杂任务的能力，但完成任务的时间成本仍明显限制了实际可用性。

表 8-1 GPT-6 在 CAD 任务中的当前能力与主要不足

测评内容	当前能力	主要不足
零件分类	能够结合几何信息与工程知识判断零件类别	细分类型判断仍不够可靠，可能将局部结构的功能误认为零件的整体用途
特征识别	能够识别孔、槽、型腔等加工特征，划分特征实例	复杂区域仍存在实例划分和类型判断错误，特别是对于嵌套、相交等复杂特征
工程图到 CAD 模型重建/生成	能够综合多视图、剖视与尺寸标注，恢复主要外形和内部结构，并完成部分多零件模型的装配	复杂结构层次与关键尺寸仍有误判，典型的有：遗漏局部特征或造成装配干涉
文本到 CAD 模型生成	能够按几何描述构造模型，依据功能需求选择设计方案，并完成建模与分析计算	复杂几何特征、孔道连通、局部过渡、零件固定及传动连接仍经常出错，功能要求与实体有效性尚不能稳定保证
图像到 CAD 模型重建/生成	能够根据多视图图像恢复主体轮廓、主要特征及零件布局	复杂曲面与连接过渡常被简化，实体连接及装配接口仍有偏差，无尺寸参照时难以确定实际尺度
B-rep 模型参数化重建	能够恢复简单和中等复杂模型的主体几何，建立参数化模型并支持部分尺寸修改	复杂曲面与内部孔道等几何的重建存在严重偏差，特征依赖关系经常不完整，修改后相关特征无法正确联动
Agentic 迭代优化	能够利用反馈补齐特征、调整尺寸与位置，降低几何误差，提高要求满足程度	仍难准确确定修改对象与范围，多轮迭代后仍可能遗留无效实体、功能缺失和参数修改失败等问题；迭代耗时较长，限制实际可用性