

英特尔 oneAPI&OpenVINO 联合黑客松竞赛题目与参赛流程

请从以下二个主题中任选一个或多个参与。

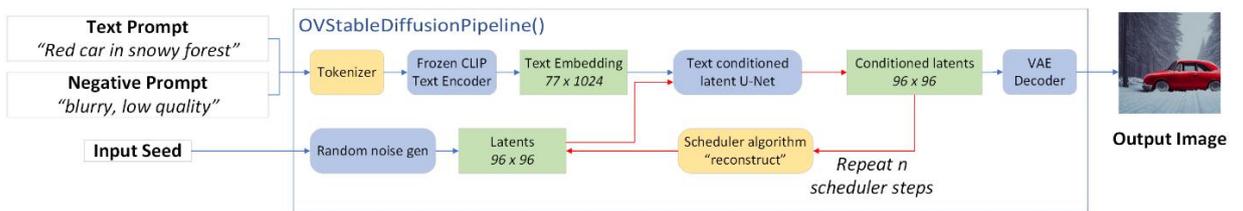
■ 赛题一：基于英特尔平台进行 Stable Diffusion 的部署优化及图片生成

本赛道旨在由参赛者使用 STABLE DIFFUSION 在目标英特尔硬件上完成部署优化及指定的图片生成工作。

操作系统： Windows 11

硬件平台： 决赛阶段入围团队统一使用大赛提供的标准硬件。

目标： 根据指定的输入输出要求，优化指定文生图模型在端侧设备上的 Pipeline 性能，在保证生图效果的情况下，降低 pipeline 端到端延迟，降低 pipeline 峰值内存占用。



模型： <https://huggingface.co/stabilityai/stable-diffusion-2-1>

模型部署及优化工具： OpenVINO (不限版本，推荐使用 GitHub 最新的开源版本)

Baseline 示例：
https://github.com/openvinotoolkit/openvino_notebooks/blob/main/notebooks/236-stable-diffusion-v2/236-stable-diffusion-v2-text-to-image.ipynb

输入输出要求：

1. Prompt 输入: "a photo of an astronaut riding a horse on mars"
2. Negative Prompt 输入: "low resolution, blurry"
3. 图片输出: 512*512, 24 Bit, PNG 格式, 参考示例:



参考优化维度：

1. pipeline 与前后处理优化
2. 模型压缩
3. Runtime Transformation (pass)
4. 模型算子 (包含使用 oneDNN 或 oneMKL 库等方式进行深入优化)

参赛流程：

1. 初赛：提交项目 Proposal，主办方将根据 Proposal 中描述的方案，评估其可行性与预计效果，并以此筛选入围决赛队伍；
2. 决赛：向入围队伍统一发放指定硬件平台，选手须围绕该硬件平台的硬件特性进行方案开发，并展示最终成果。

决赛评价标准：

决赛将从以下 3 个维度对作品进行评分，每个维度占 10 分，单项第一名得满分 10 分，第二名得 9 分，以此类推。此外，考虑到生成图像的真实性，参赛选手需保证最后生图得主观效果，即该 pipeline 所生成的图像，需要与主题范围内 Prompt 所描述的场景大体一致，且真实，反之 2, 3 评价标准将各积 1 分。

1. 方案完整性，可行性与创新性；
2. 在相同生成参数 (除随机种子外)，及相同 prompt 输入情况下，随机生成 10 张不同的图片 (每张图片需包含 20 次采样迭代)，统计 10 张图片生成过程中的端到端总延迟 (文本输入-最后一张 RGB 图像输出) 越低越好，2 个作品相差 $\pm 5\%$ 以内将积同分；
3. 在相同生成参数 (除随机种子外)，及相同 prompt 输入情况下，随机生成 10 张不同的图片 (每张图片需包含 20 次采样迭代)，统计过程中的内存占用峰值，越低越好，2 个作品相差 $\pm 5\%$ 以内将积同分。

加分项：

若前三项总分相同，我们将根据 CPU 的使用率峰值进行额外评估，使用率越低，分数越高

其他限制：

1. 模型优化必须在指定模型结构基础上进行优化，除量化与 LoRa 等微调技术外，不得改变模型原始结构；
2. 决赛参赛作品需要在指定硬件平台上以及统一的操作系统上运行，从而获得评价指标。

■ 赛题二：基于英特尔® oneAPI 的开放创新

问题描述：

- 创意可以是来自任何领域的任何设想，没有限制，唯一的要求是，您需要在您的方案中使用到下列两类技术工具中的至少一类，为特定问题构建解决方案。
- 第一类：[英特尔® oneAPI AI 分析工具套件](#) (INTEL® AI ANALYTICS TOOLKITS) 中的一项或多项产品组件及库
- 第二类：[英特尔® oneAPI 基础工具套件](#) (INTEL® ONEAPI BASE TOOLKIT) 中的

C++/SYCL 直接编程语言及适用的 API 函数库，或使用 DPCT 工具完成可运行的 SYCL 代码迁移

评选规则：

1. 代码质量

- 代码易于理解且可重现
- 代码经过良好测试，功能无错误
- 代码有详细记录

2. 整体可用性

- 开发人员适当利用和使用英特尔® AI 分析工具包或英特尔® oneAPI 基础工具套件
- 解决方案的完成度
- 解决方案的输出是否与所描述的问题相关度和效果
- 解决方案应在技术上可行，提出的解决方案应在实施方面要切合实际
- 解决方案的易用程度，应该是用户友好的、易于使用且易于理解的

5. 价值影响

- 解决方案是否提出了一个有潜力的商业模式
- 该解决方案应该是可扩展的、可持续性
- 该解决方案是否会对行业产生价值影响和颠覆