

PR #1760 完整报告

THUDM/slime

[Multimodal] Add Multimodal OPD support

合并时间: 2026-03-27 15:44

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1760>

执行摘要

PR 1760 为 on-policy distillation 添加了多模态输入支持，通过处理 `sample` 中的图像数据并添加到请求负载，解决了 VLM 在蒸馏过程中无法传递图像信息的问题。这是一个小型但关键的改进，扩展了 SLIME 对视觉语言模型的支持，确保图像能正确集成到蒸馏流程中。

功能与动机

动机源于 issue #1758，用户询问 SLIME 是否支持 VLMs 的 on-policy distillation，因为当前 `sample.tokens` 不包含图像信息，导致 SGLang teacher server 无法处理图像，从而无法进行蒸馏。此 PR 旨在修复此问题，使多模态模型能够顺利进行蒸馏，引用 issue 中的表述："Does slime currently support OPD for VLMs?" 和用户确认的兼容性问题。

实现拆解

修改了 `slime/rollout/on_policy_distillation.py` 文件中的 `reward_func` 函数。关键改动点：

- 导入 `encode_image_for_rollout_engine` 函数以处理图像编码。
- 在函数中添加条件代码块：

```
python if sample.multimodal_inputs and
sample.multimodal_inputs.get("images"): image_data =
sample.multimodal_inputs["images"] payload["image_data"] =
[encode_image_for_rollout_engine(image) for image in image_data]
```
- 这确保了当 `sample` 包含图像时，将编码后的图像数据作为 `image_data` 列表添加到 `payload` 中，传递给奖励模型。

评论区精华

在 issue 评论中，主要讨论围绕如何正确传递多模态数据：

- GentleZhu指出: `"text": sample.prompt + sample.response, works but existing 'input_ids': sample.response won't because SGLang teacher server won't know images are being passed. It requires the vision marker."`
- coding-famer回应: 建议使用 `"input_ids": sample.tokens`，因为 `sample.tokens` 包含图像 token，并解释 `input_ids` 会在 SGLang 内部转换回文本。讨论强调了正确集成图像信息的重要性，但 PR 采用添加 `image_data` 字段的方式，可能是一种折中方案。

风险与影响

风险:

1. 依赖风险: 新增导入 `encode_image_for_rollout_engine` 函数, 需确保该函数稳定且与图像格式兼容。
2. 数据结构风险: 条件检查依赖 `sample.multimodal_inputs` 结构, 若结构变更或未正确定义, 可能导致错误。
3. 测试覆盖不足: 缺少专门测试, 可能在多模态场景下引入回归。

影响:

- 用户: 现在可使用 VLMs 进行 on-policy distillation, 提升模型训练灵活性。
- 系统: 增加图像处理逻辑, 需确保与 SGLang server 接口兼容, 可能影响请求性能。
- 团队: 可能需要更新文档或添加多模态示例, 以指导用户使用。

关联脉络

与历史 PR 1745 "feat: GLM4V multimodal support improvements" 直接相关, 后者改进了图像处理和训练兼容性, 而此 PR 专注于 on-policy distillation 的多模态支持。两者共同推动 SLIME 在多模态领域的演进, 形成从基础处理到高级蒸馏的功能链条。关联 issue #1758 揭示了用户需求, 驱动了此改进的实现。