

# PR #5874 完整报告

verl-project/verl

[megatron, cfg] feat: add Qwen3.5-122B Megatron launch script

合并时间: 2026-04-03 14:20

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5874>

## 执行摘要

- 一句话: 新增 Qwen3.5-122B Megatron 启动脚本, 支持 32 GPU 大规模 GRPO 训练。
- 推荐动作: 该 PR 对于需要运行 Qwen3.5-122B 或类似大规模模型的工程师值得参考, 特别是关注 Megatron 并行配置 (如 TP、PP、CP 设置) 和模型特定限制 (如 GDN 注意力格式)。建议精读脚本中的配置注释, 以了解架构权衡和未来优化方向。

## 功能与动机

根据 PR body, 动机是成功运行 Qwen3.5-122B GRPO 训练并提供示例配置, 具体表述为: 'This PR successfully runs Qwen3.5-122B GRPO on verl with mbridge support for Qwen3.5 using 32 H20 GPUs.' 旨在帮助用户复现类似大规模训练场景。

## 实现拆解

实现方案仅为添加一个 bash 脚本文件 'examples/grpo\_trainer/run\_qwen3\_5-122b-a10b-megatron.sh'。脚本包含环境变量设置 (如 CUDA\_VISIBLE\_DEVICES)、数据集下载命令、模型路径配置、并行参数 (TP=2, PP=2, CP=1, EP=8, ETP=1, GEN\_TP=8) 以及训练超参数 (如 batch size、prompt length)。关键改动点是针对 Qwen3.5 架构的 Gated Delta Net 线性注意力不支持 packed sequences, 因此设置 use\_remove\_padding=False 和 use\_dynamic\_bsz=False 以使用 bshd 格式, 并注释了未来 Megatron-LM 支持 THD 格式时的优化方向。

关键文件:

- examples/grpo\_trainer/run\_qwen3\_5-122b-a10b-megatron.sh (模块 examples/grpo\_trainer): 新增 Qwen3.5-122B Megatron 启动脚本, 提供完整训练配置示例, 包括环境设置、并行策略和超参数, 是 PR 的唯一变更文件。

关键符号: 未识别

## 评论区精华

review 中由 gemini-code-assist[bot] 提出两个核心讨论: 1. 命令兼容性问题: 'hf' 命令应改为 'huggingface-cli' 以确保脚本可移植性, 但提交历史未明确显示是否修复; 2. 性能优化问题: filter\_overlong\_prompts\_workers 初始设置为 256 过高, 建议改为 64 以避免资源耗尽, 该问题在第二次提交中修复。结论是第二个问题已解决, 第一个问题仍存在潜在风险。

- 命令兼容性问题 (correctness): 建议修改, 但提交历史未显示是否修复, 仍存在潜在兼容性风险。
- worker 数量设置优化 (performance): 在第二次提交中修复, 改为 64, 已解决性能问题。

## 风险与影响

- 风险: 技术风险包括: 1. 命令兼容性: 脚本使用非标准 'hf' 命令, 可能在某些环境失败, 影响用户运行; 2. 资源设置: 初始 worker 数量过高可能导致性能下降或资源耗尽, 但已修复; 3. 环境依赖: 脚本依赖特定 Docker 镜像 (verlari/verl:vllm017.latest) 和包版本 (如 Megatron-LM==0.16.0), 不兼容环境可能引发运行错误; 4. 配置复杂性: 并行参数和模型特定设置 (如 use\_remove\_padding=False) 需要用户理解架构限制, 错误配置可能影响训练效率。
- 影响: 影响范围主要针对用户示例和配置参考: 为用户提供运行 Qwen3.5-122B 的实用脚本, 有助于加速模型部署和实验复现。对系统核心代码无直接影响, 但可能促进更多用户采用 Megatron 后端进行大规模训练。影响程度中等, 仅限于示例目录和用户配置层面。
- 风险标记: 命令兼容性风险, 资源配置过高, 依赖特定环境

## 关联脉络

- PR #5848 [cfg] refactor: unify ppo\_trainer and ppo\_megatron\_trainer config: 统一训练器配置, 为本脚本中的 Megatron 配置提供基础, 涉及 config 模块。
- PR #5864 [fully\_async] chore: Update fully async dapo qwen3-30b npu script: 类似脚本更新示例, 涉及修改训练脚本以优化超参数和资源设置。
- PR #5057 [megatron] feat: support dynamic CP: Megatron 动态上下文并行支持, 与本脚本中的 CP 配置相关, 涉及性能优化。