

PR #7056 完整报告

PaddlePaddle/FastDeploy

[CI] Align with Paddle layer_norm kernel update

合并时间: 2026-03-27 22:58

原文链接: <http://prhub.com.cn/PaddlePaddle/FastDeploy/pull/7056>

执行摘要

本 PR 更新了 FastDeploy 的 CI 测试基线，以适应上游 Paddle layer_norm 内核的变更，确保测试通过并维持 CI 稳定性。这是一个低风险的维护性变更，仅涉及测试文件修改，不影响核心功能。

功能与动机

上游 Paddle PR #78321 引入了 accuracy-compatible 的 layer_norm 内核，改变了数值计算行为（如 epsilon 精度和均值 / 方差计算），导致 FastDeploy CI 中 Qwen3VLMoe 测试出现精度漂移。为避免误报 CI 失败，需要调整测试基线以对齐新行为，保持与上游的一致性。

实现拆解

修改涉及两个文件：

- tests/cov_pytest.ini: 从忽略列表中移除 test_Qwen3VLMoe_serving.py（删除行 `--ignore=tests/e2e/test_Qwen3VLMoe_serving.py`），使该测试重新纳入覆盖检查。
- tests/e2e/test_Qwen3VLMoe_serving.py: 在 test_consistency_between_runs 函数中，将预期输出字符串从 "根据您提供的视频片段，我们可以观" 更新为 "根据您提供的视频帧，我们可以观"，以反映 layer_norm 变更后的模型输出。

评论区精华

无 review 讨论；在 issue 评论中，作者发布了跳过多个 CI 任务的命令（如 `/skip-ci ci_iluvatar`），可能用于加速合并。Codecov 报告显示代码覆盖良好，无覆盖问题。

风险与影响

风险较低，仅影响测试基线，不涉及核心功能，因此无直接回归、性能或安全风险。但需注意对外部上游 Paddle 的依赖，未来类似变更可能需再次调整测试基线，存在外部变更同步风险。对用户无直接影响，但有助于保持 CI 稳定和团队开发效率。

关联脉络

从近期历史 PR 分析中，未发现直接相关 PR；但此变更反映了 FastDeploy 与上游 Paddle 保持对齐的持续维护策略，类似于其他 CI 或测试更新，凸显了在依赖外部库时维护测试基线的重要

要性。