

PR #2607 完整报告

radixark/miles

docs: fix diffusion docs

合并时间: 2026-08-18 19:37

原文链接: <http://prhub.com.cn/radixark/miles/pull/2607>

执行摘要

继 PR#2583 将 diffusion 文档移植到主站后, 本 PR 对这批文档做内容校准与导航补齐: 将 Cosmos3 页面从覆盖整个 MoT omni 家族的通用说明收敛为仅覆盖有 canonical recipe 与训练曲线验证的 Cosmos3-Nano; 在顶层模型表新增 SD3.5、Qwen-Image、Wan2.2-T2V-A14B、LTX-2.3、Cosmos3-Nano 五个入口; 并在 diffusion 首页与顶层模型表之间建立双向链接。全部为纯文档改动 (3 个文件、+18/-17), 无代码与测试变更, 维护者 Zhichenzzz 已直接批准。

功能与动机

PR body 的两条要点直接给出动机: 一是将已验证的 diffusion 模型家族暴露到顶层模型表, 解决用户从主文档难以发现这些 recipe 的问题; 二是把嵌入的 Cosmos3 文档限制到唯一有 canonical recipe 与训练曲线支撑的 Nano 变体。换言之, 本次修正同时解决“可发现性不足”和“声明超出验证证据”两个问题——原页面宣称整个 Cosmos3 家族适用, 但实际被验证的只有 Nano, Edge 与 Super 均标为 Untested, 继续保留家族式表述会误导用户选型。

实现拆解

1. 顶层模型表补齐 diffusion 入口 (docs/index.md): 在 Supported models 大表末尾新增 Stable Diffusion、Qwen-Image、Wan、LTX、Cosmos 五行, 分别链接到 /diffusion/models/ 下的 sd3、qwen-image、wan、ltx、cosmos 页面; 同时把原先整段“See Models”指引拆分为 LLM 索引与 Diffusion 索引两条路径, 避免两类文档相互淹没。
2. 收敛 Cosmos3 文档范围 (docs/diffusion/models/cosmos/cosmos3.md): 将页面 description 从面向 Cosmos3 MoT omni 家族 (UND + GEN towers) 的通用说明改为仅描述 Cosmos3-Nano 的 Flow-GRPO + PickScore recipe; 删除原表格中 Edge (4 B) 与 Super (64 B) 两行标为 Untested 的条目, 并移除“All sizes resolve to the same family config”的家族式表述, 使声明与验证证据严格对齐。
3. 同步 diffusion 首页并打通导航回路 (docs/diffusion/index.md): 核心特性列表中的“the Cosmos3 MoT omni family”改为“Cosmos3-Nano”; Supported models 段补充“Validated models also appear in the Miles model list”并链接回顶层模型表锚点, 与第 1 步形成双向可达。
4. 校验与测试配套: 本次无源码、测试、配置或 schema 改动; 作者执行了 git diff --check 与文档诊断, 确认无错误, 并按 docs-only 豁免不运行 pre-commit 与 pytest。

顶层模型表新增 diffusion 入口 (docs/index.md)

```
<!-- 顶层 Supported models 大表末尾新增的 Diffusion 家族入口：
    每条链接指向 /diffusion 下独立 recipe 指南，与 diffusion 子页形成双向可达 -->
| **Stable Diffusion** | [SD3.5](/diffusion/models/sd3/sd3) |
| **Qwen-Image** | [Qwen-Image](/diffusion/models/qwen-image/qwen-image) |
| **Wan** | [Wan2.2-T2V-A14B](/diffusion/models/wan/wan2-2) |
| **LTX** | [LTX-2.3](/diffusion/models/ltx/ltx2) |
| **Cosmos** | [Cosmos3-Nano](/diffusion/models/cosmos/cosmos3) |
```

Cosmos3 支持范围收敛 (docs/diffusion/models/cosmos/cosmos3.md)

```
<!-- 修订后的 Supported variants 部分：范围收敛到唯一有 canonical recipe 与训练曲线
    支撑的 Cosmos3-Nano；Edge 与 Super 因缺少验证证据不再列出，避免未验证声明误导用户 -->
| Size | Composition | HF checkpoints | Status |
| --- | --- | --- | --- |
| Nano | 16 B (8 B UND + 8 B GEN) | [Cosmos3-Nano](https://huggingface.co/nvidia/Cosmos3-Nano)<br/>[Cosmos3-Nano-Policy-DROID](https://huggingface.co/nvidia/Cosmos3-Nano-Policy-DROID) | **Validated** — canonical recipes |
```

评论区精华

本 PR 的 review 没有实质性技术交锋，属于低争议、范围清晰的文档修正：

- claude[bot] 仅提示这是 fork PR，自动 review 被禁用，维护者可评论 @claude review 触发一次性审核，该提议未被采纳。
- 维护者 Zhichenzzz 直接批准且未留评论，说明改动与既有文档方向一致，无悬而未决的疑虑。

风险与影响

- 链接一致性：docs/index.md 新增的 /diffusion/models/sd3/sd3、/diffusion/models/cosmos/cosmos3 等路径必须与 docs/docs.json 路由一致，否则生成站点会出现 404；文档诊断已通过，风险较低。
- 表述误读：删除 Edge/Super 状态行后，Cosmos3 页面可能被误读为“仅支持 Nano”；但原表本就将两者标为 Untested，收敛属于保守表述而非支持范围缩小。
- 验证级别同步：docs/diffusion/index.md 的 recipe 验证级别标签需随后续验证进度同步更新，否则与顶层表入口会出现信息不同步。
- 影响范围：对系统无运行时代码影响；对用户可发现性有正面提升；对团队确立了“文档只声明有验证证据支撑的模型”的写作纪律。

关联脉络

本 PR 是 PR#2583（移植 miles_diffusion 全套指南到主站）的后续修正，把 Cosmos3 页面从家族级声明收敛为 Nano 验证级声明；与 PR#2599（README 与 docs 首页补充 diffusion 后训练指引）同属 diffusion 文档入口建设；与 PR#2606（dashboard 文档回退与重定向）一起构成近期文档结构与可发现性的持续优化。整体趋势是：主站文档体系正从“搬运内容”走向“声明与验证证据严格对齐 + 顶层目录与子页双向导航”的精细化阶段。