

DeepSeek-OCR 架构深度解析

探索 DeepSeek-OCR 如何通过创新的编码器-解码器架构,实现文本的光学压缩与高保真解码

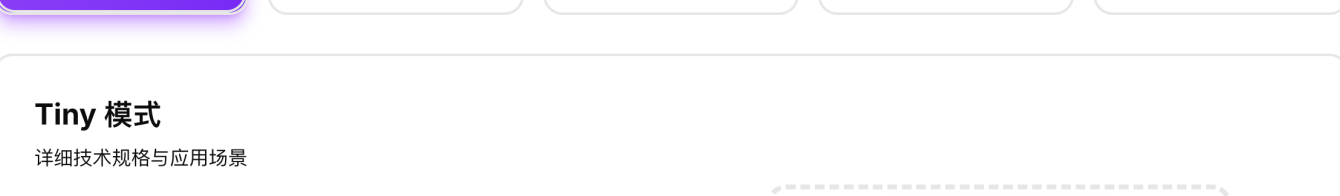
DeepSeek-OCR 整体架构

采用统一的编码器-解码器架构, 通过创新的光学压缩技术实现文本信息的高效处理



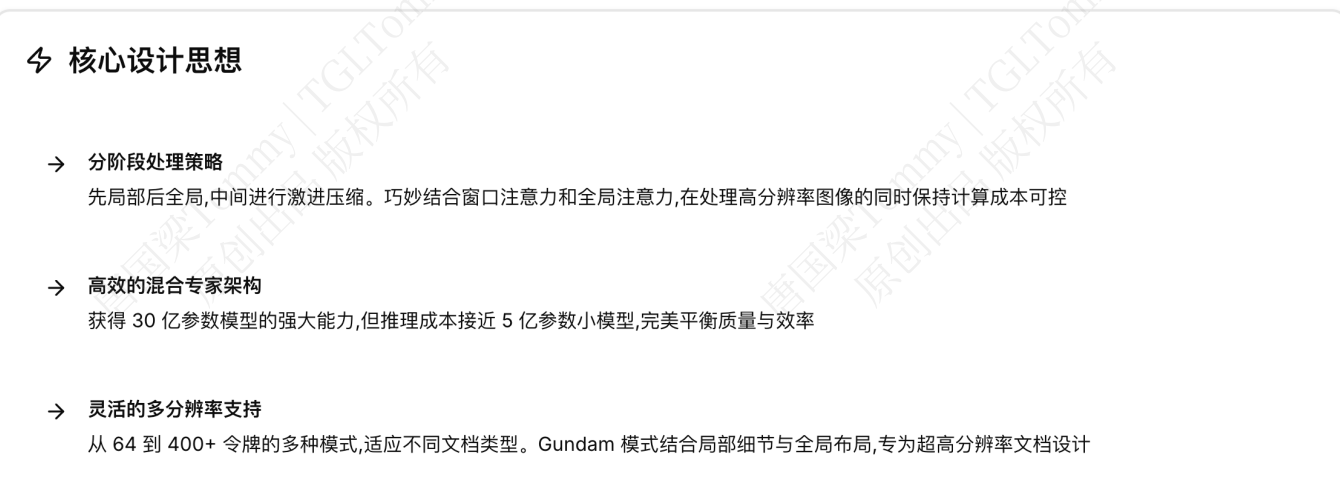
DeepEncoder 四步工作流程

DeepEncoder 采用“先粗后精、层层递进”的策略,通过四个关键步骤实现高效的图像压缩。从4096个图像块压缩到256个视觉令牌,为后续解码器提供高度浓缩的视觉表示。



MoE 解码器工作原理

解码器的核心任务是“解压缩”——它接收由DeepEncoder生成的、高度浓缩的视觉令牌,并将其“翻译”回详细、可读的文本。这是整个系统的第二个核心组件,负责将压缩的视觉表示还原为原始文本内容。



课程官网: TGLTommy.com (国内访问需科学上网)

AI进阶学习

B站课堂: <https://space.bilibili.com/3546748815411393/pugv>

官方网站: <https://www.tgltommy.com/> (国内访问需科学上网)

如果你正在关注大模型 Agent、强化学习后训练、RLHF、DPO、GRPO、RLVR 等前沿方向, 欢迎关注我最新上线的精品课程: 《大模型 Agent 强化学习实战: 从后训练、奖励设计到工业级对齐系统》

这门课程围绕当前大模型 Agent 强化学习的核心技术路线展开, 不只是介绍零散算法和论文概念, 而是希望帮助学习者建立一套完整的技术地图: 从基础原理、论文脉络、关键算法, 到开源项目源码解析与工业级系统实践, 逐步理解 Agent RL 为什么重要、如何演进, 以及在真实应用中如何落地。

课程配套专属电子书知识库《Agent RL 学习宝典》, 包含学习路线、论文资料、项目说明、源码阅读辅助资料和后续持续更新内容, 方便长期查阅和跟进前沿技术。

Agent RL 变化很快, 但越是变化快, 越需要一套系统的学习框架。如果你希望系统进入 Agent RL 方向, 而不是停留在碎片化了解阶段, 这门课会是一个很好的起点。

课程官网: <https://www.tgltommy.com/p/agent-rl> (国内访问需科学上网)

B站课堂: <https://www.bilibili.com/cheese/play/ss842375604>

更多课程详情, 扫描下方二维码咨询课程助理:



唐国梁Tommy 原创出品 | © 2026 倚智未来 (深圳) 科技有限公司。保留所有权利。

课程官网: TGLTommy.com (国内访问需科学上网)



唐国梁Tommy 原创出品 | © 2026 倚智未来 (深圳) 科技有限公司。保留所有权利。