Artificial Intelligence Newsletter

人工智能前沿快报

2025 年第 10 期
总第 24 期

学会官方公众号

广东省图象图形学会
GDSIG

学术交流、科普宣传、科技成果鉴定

目 录

学 术 动 态

目 录

技 术 动 态

资 源 分 享

鸣 谢 支 持

INTRODUCTION
导读

在人工智能技术飞速发展的时代背景下，我们将不定期梳理人工智能领域的部分前沿学术与技术动态，并以简报的形式分享给读者，以帮助关注此领域的人士了解最新的学术前沿发展与产业技术动态，促进学术交流和技术繁荣。本期摘选了近期全球人工智能领域的 22 篇最新学术动态、7 篇技术动态、以及 1 个资源推荐给广大读者。

ACADEMIC TRENDS
学术动态

● 论文一

Sa2VA: Marrying SAM2 with LLaVA for Dense Grounded Understanding of Images and Videos

日期

2025-01-07

Sa2VA: Marrying SAM2 with LLaVA for Dense Grounded Understanding of Images and Videos

Haobo Yuan[1*], Xiangtai Li[2* †], Tao Zhang[2,3*], Zilong Huang[2], Shilin Xu[4], Shunping Ji[3], Yunhai Tong[4], Lu Qi[2], Jiashi Feng[2], Ming-Hsuan Yang[1]
[1]UC Merced [2]Bytedance Seed [3]Wuhan University [4]Peking University

内容摘要

本文介绍了一种名为 Sa2VA 的新颖框架，它首次将一个顶尖的视频分割模型（SAM-2）与一个强大的视觉语言大模型（LLaVA）整合进一个单一的、可端到端训练的系统中。其目标是实现对图像和视频的“密集扎根的理解”（dense grounded understanding），即模型不仅能就视觉内容进行对话，还能精确地分割和追踪对话中提到的物体。该模型的核心架构采用了一种解耦设计：一个类 LLaVA 模型处理文本、图像和视频输入，以生成语言回复和一个特殊的 [SEG] 令牌（token）。这个 [SEG] 令牌的隐藏状态随后被用作一个学习到的提示（prompt），来指导 SAM-2 的解码器生成相应的分割掩码（mask）。这种方法使得模型能够通过统一的、一次性指令微调（one-shot instruction tuning）来处理多种任务，包括指代性图像/视频分割、图像/视频聊天以及扎根的字幕生成。为了解决缺乏具挑战性训练数据的问题，作者还推出了 Ref-SAV，这是一个用于复杂场景下指代性视频对象分割的大规模自动标注数据集。

Figure 1. **Illustration of capabilities of our proposed Sa2VA.** (a). Given a video, Sa2VA is able to segment the referred object and understand the whole scene. (b).Sa2VA supports image conversation, video conversation, image referring segmentation, video referring segmentation, and grounded caption generation with single-shot instruction-tuning. (c).Sa2VA achieves strong results on multiple images, video referring segmentation, and chat benchmarks compared with existing MLLMs, such as GLaMM [66] and OMG-LLaVA [99].

主要亮点

- **首个用于密集扎根图像和视频理解的统一模型：**Sa2VA 是第一个将视频分割基础模型（SAM-2）与视觉语言大模型（LLaVA）“联姻”成一个单一端到端系统的框架。这种独特的结合使其能够同时执行对话任务（聊天）和密集的感知任务（分割），并且同时适用于静态图像和动态视频，这是以往模型所不具备的能力。
- **通过共享[SEG]令牌实现的优雅解耦架构：**该模型使用了一种简单而有效的通信机制，即由大语言模型生成一个特殊的 [SEG] 令牌，其隐藏状态用于提示 SAM-2 解码器。这种解耦的、“即插即用”的设计使得框架能够保留两个基础模型的强大预训练知识，并且便于未来升级使用更新的 MLLM 或分割模型。
- **推出 Ref-SAV 基准和数据集：**为了推动视频扎根理解领域的发展，该论文推出了 Ref-SAV，这是一个全新的、富有挑战性的指代性视频对象分割基准。该数据集通过自动化标注流程创建，包含了超过 7.2 万条在复杂视频中的对象描述，这些视频具有严重的遮挡和长文本描述，为研究社区提供了宝贵的新资源。

相关链接

Paper: https://arxiv.org/abs/2501.04001
GitHub: https://github.com/bytedance/Sa2VA
Model： https://huggingface.co/collections/ByteDance/sa2va-model-zoo

● 论文二

Increasing alignment of large language models with language processing in the human brain

日期

2025-09-16（Nature computational science 录用论文）

nature computational science

Article

https://doi.org/10.1038/s43588-025-00863-0

Increasing alignment of large language models with language processing in the human brain

Received: 20 September 2024

Accepted: 7 August 2025

Published online: 16 September 2025

Check for updates

Changjiang Gao[1,2,4], Zhengwu Ma[1,4], Jiajun Chen[2], Ping Li[3], Shujian Huang[2,5] & Jixing Li[1,4,5]

Transformer-based large language models (LLMs) have considerably advanced our understanding of how meaning is represented in the human brain; however, the validity of increasingly large LLMs is being questioned due to their extensive training data and their ability to access context thousands of words long. In this study we investigated whether instruction tuning—another core technique in recent LLMs that goes beyond mere scaling—can enhance models' ability to capture linguistic information in the human brain. We compared base and instruction-tuned LLMs of varying sizes against human behavioral and brain activity measured with eye-tracking and functional magnetic resonance imaging during naturalistic reading. We show that simply making LLMs larger leads to a closer match with the human brain than fine-tuning them with instructions. These finding have substantial implications for understanding the cognitive plausibility of LLMs and their role in studying naturalistic language comprehension.

内容提要

本文系统性地研究了大型语言模型（LLM）与人脑语言处理的对齐问题，旨在探究是扩大模型规模（scaling）还是进行指令微调（instruction tuning）更能提升模型的认知真实性。尽管 LLM 在模拟人类语言处理方面潜力巨大，但其日益增长的规模和超人类的训练数据使其作为认知模型的有效性引发了广泛争论。为解决此问题，本研究将不同规模的基础 LLM 和经过指令微调的 LLM 的内部自注意力机制，分别与人类在自然阅读任务中的眼动行为数据和功能性磁共振成像（fMRI）脑活动数据进行回归分析。研究结果明确指出，仅仅扩大模型规模（从 7 亿到 650 亿参数）比进行指令微调能带来与人脑活动更紧密的对齐，即更大的模型能更好地预测人类的阅读行为与神经反应，而指令微调并未显示出优势。这一发现对于理解 LLM 的认知合理性以及其在自然语言理解研究中的应用具有重要的启示意义。

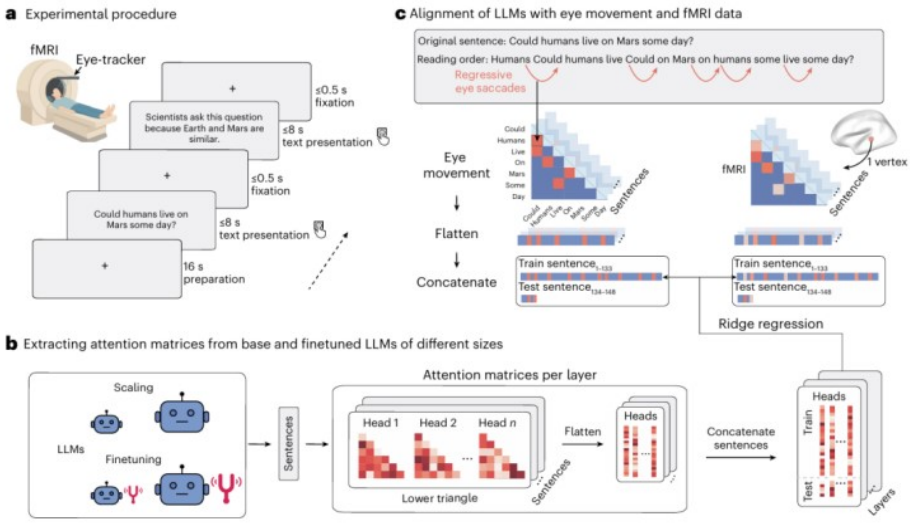


Fig. 1 | Experimental procedure of the dataset and the analyses pipeline. **a**, Experiment procedure. Participants read five English articles sentence-by-sentence while inside the fMRI scanner with concurrent eye-tracking. **b**, LLMs of different sizes with and without fine-tuning are employed in the study. **c**, Analysis pipeline. The attention matrices of each layer of the LLMs for each sentence in the experimental stimuli were averaged over attention heads and aligned with eye movement and fMRI activity patterns for each sentence using ridge regression.

主要亮点

- **规模与微调的系统性比较：**首次系统性地对比了扩大模型规模（scaling）和进行指令微调（instruction tuning）这两种主流技术，旨在明确哪种方式能更有效地提升大型语言模型与人脑语言处理机制的对齐程度。
- **扩大规模对脑对齐的关键作用：**研究明确发现，模型的规模越大，其内部表征与人类在自然阅读时的眼动行为和脑活动模式的对齐程度就越高，证实了“规模法则”（scaling law）在认知真实性上也同样适用，而指令微调对此并无显著增益。
- **多模态数据的对齐分析：**创新性地将模型的自注意力矩阵与人类的两种生理数据——眼动追溯（行为）和功能性磁共振成像（神经）——进行对齐分析，为评估模型的认知合理性提供了全面的跨层级证据。

相关链接

Paper: https://doi.org/10.1038/s43588-025-00863-0

● 论文三

Self-Adapting Language Models

日期

2025-09-18

Self-Adapting Language Models

Adam Zweiger* **Jyothish Pari*†** **Han Guo** **Ekin Akyürek** **Yoon Kim** **Pulkit Agrawal†**
Massachusetts Institute of Technology
{adamz, jyop, hanguo, akyurek, yoonkim, pulkitag}@mit.edu

内容提要

本文提出了自适应语言模型（Self-Adapting Language Models, SEAL），一个旨在解决大型语言模型（LLMs）在预训练后本质上是静态的、缺乏适应新任务或新知识能力的框架。该框架的核心思想是让语言模型通过生成自己的微调数据和更新指令（称为“自编辑”，self-edit）来实现自我调整。具体来说，当面对新输入时，模型会生成可能包含信息重组、优化超参数指定或工具调用等内容的“自编辑”，并通过监督式微调 (SFT) 将这些编辑内容应用到自身权重上，从而实现持久的适应。为了训练模型生成有效的“自编辑”，文章采用了一个强化学习循环，将模型更新后在下游任务上的性能提升作为奖励信号。实验证明，无论是在知识整合还是在少样本泛化任务上，SEAL 都展现了其有效性，为实现能够响应新数据并进行自我指导式学习的语言模型迈出了有前景的一步。

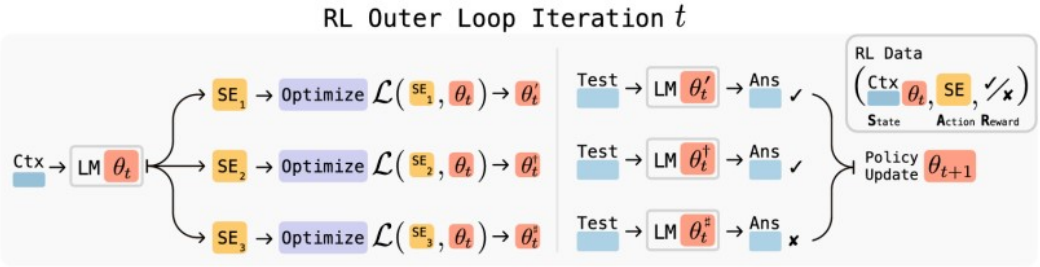


Figure 1: **Overview of SEAL.** In each RL outer loop iteration, the model generates candidate self-edits (SE)—directives on how to update the weights—applies updates, evaluates performance on a downstream task, and uses the resulting rewards to improve the self-edit generation policy.

主要亮点

- **自编辑与自更新机制：**提出 SEAL 框架，使语言模型能够生成自己的微调数据和更新指令（即“自编辑”），并通过监督式微调（SFT）实现对自身模型权重的持久性更新，从而摆脱静态模型的限制。
- **强化学习优化循环：**创新性地采用一个强化学习（RL）外循环来优化“自编辑”的生成策略。模型的更新效果（即在下游任务上的性能提升）直接作为奖励信号，从而引导模型学会如何生成最有效的训练数据来改进自己。
- **统一的自适应参数化：**与依赖外部适配器模块或超网络的方法不同，SEAL 直接利用模型自身的生成能力来参数化和控制整个适应过程。模型通过生成自然语言指令来定义训练数据和优化策略，具有更高的通用性和灵活性。

相关链接

Paper: https://arxiv.org/abs/2506.10943
GitHub: https://jyopari.github.io/posts/seal

● 论文四

ReasoningBank: Scaling Agent Self-Evolving with Reasoning Memory

日期

2025-09-29

REASONINGBANK: Scaling Agent Self-Evolving with Reasoning Memory

Siru Ouyang[1*], Jun Yan[2], I-Hung Hsu[2], Yanfei Chen[2], Ke Jiang[2], Zifeng Wang[2], Rujun Han[2], Long T. Le[2], Samira Daruki[2], Xiangru Tang[3], Vishy Tirumalashetty[2], George Lee[2], Mahsan Rofouei[4], Hangfei Lin[4], Jiawei Han[1], Chen-Yu Lee[2] and Tomas Pfister[2]
[1]University of Illinois Urbana-Champaign, [2]Google Cloud AI Research, [3]Yale University, [4]Google Cloud AI

内容提要

本文提出了 ReasoningBank，一个旨在让大型语言模型（LLM）代理能够从过往经验中持续学习并自我进化的新颖记忆框架。针对现有代理在面对连续任务时无法有效利用历史交互、重复犯错的问题，ReasoningBank 通过自我判断，从成功和失败两种经验中提炼出可泛化的、高层次的推理策略并存入记忆库。当代理遇到新任务时，它会检索相关策略以指导决策，并在任务完成后将新经验的感悟整合回库，形成一个闭环学习系统。在此基础上，文章进一步引入了内存感知的测试时扩展（MaTTS）技术，通过在单个任务上增加计算投入以产生多样化的探索经验，为 ReasoningBank 提供了丰富的对比信号来生成更高质量的记忆，从而在记忆与测试时扩展之间建立了强大的协同效应。通过在网页浏览和软件工程等基准测试上的广泛实验，该研究证明了 ReasoningBank 不仅提升了代理的效率和成功率，还确立了以记忆驱动的经验扩展作为一种新的能力提升维度，使代理能够自然地涌现出更复杂的行为。

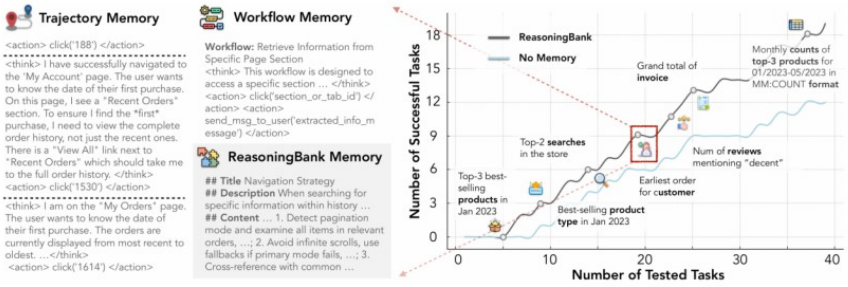


Figure 1 | REASONINGBANK induces reusable reasoning strategies, making memory items more transferrable for future use. This enables agents to continuously evolve and achieve higher accumulative success rates than the "No Memory" baseline on the WebArena-Admin subset.

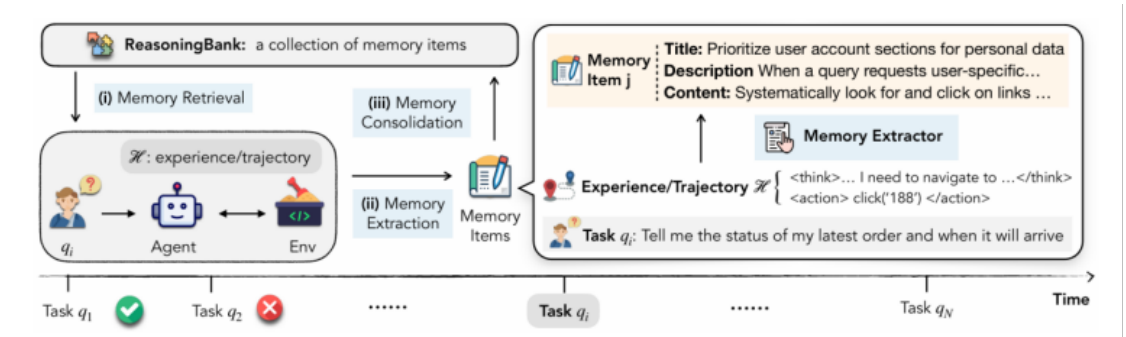


Figure 2 | Overview of REASONINGBANK. Experiences are distilled into structured memory items with a title, description, and content. For each new task, the agent retrieves relevant items to interact with the environment, and constructs new ones from both successful and failed trajectories. These items are then consolidated into REASONINGBANK, forming a closed-loop memory process.

主要亮点

- **双重经验提炼的记忆框架：**提出 ReasoningBank 框架，它不仅记录成功的经验，更关键的是能从失败的交互中提炼出可泛化的推理策略和教训，使代理能够避免重复犯错，超越了仅依赖成功轨迹的传统记忆方法。
- **记忆与扩展的协同效应：**引入内存感知的测试时扩展（MaTTS）技术，通过在单个任务上进行深度探索，为 ReasoningBank 提供丰富的成功与失败对比信号以生成更高质量的记忆。这种高质量记忆反过来又指导扩展过程，实现了记忆与计算扩展之间的强大协同效应。
- **实现代理的自我进化与能力涌现：**构建了一个闭环的自我进化系统，使代理能够在其生命周期内持续从任务流中学习。实验证明，该方法不仅显著提升了任务成功率和效率，还促使代理自然涌现出从简单到复杂的推理策略，真正实现了能力的持续成长。

相关链接

Paper: https://arxiv.org/abs/2509.25140

● 论文五

Interpreting Language Models Through Concept Descriptions: A Survey

日期

2025-10-01（录用于 BlackboxNLP Workshop）

Interpreting Language Models Through Concept Descriptions: A Survey

Nils Feldhus[*1,2] **Laura Kopf**[*1,2]
[1]BIFOLD – Berlin Institute for the Foundations of Learning and Data
[2]Technische Universität Berlin
{feldhus,kopf}@tu-berlin.de
*Equal contribution

内容提要

本文是一篇综述，系统性地介绍了通过“概念描述”来解释大型语言模型（LLM）内部机制这一新兴领域。理解神经网络的决策过程是机械可解释性的核心目标，而近期研究开始利用强大的生成模型，为目标模型的内部组件（如神经元、注意力头）和模型抽象（如稀疏自编码器 SAE 特征）自动生成开放词汇的自然语言描述，以揭示其功能。本文对这一快速发展的领域进行了首次全面梳理，详细阐述了生成这些概念描述的关键方法、不断演进的自动化与人工评估指标体系（涵盖预测模拟、因果干预等），以及支撑该研究的核心数据集。综述发现，当前研究正从解释单个多义性神经元转向分析更纯粹的模型抽象（如 SAE 特征），并且对描述的评估越来越强调严格的因果验证。通过勾勒该领域的现状、趋势与挑战，本文为未来研究提供了一份清晰的路线图，旨在最终提升模型的透明度。

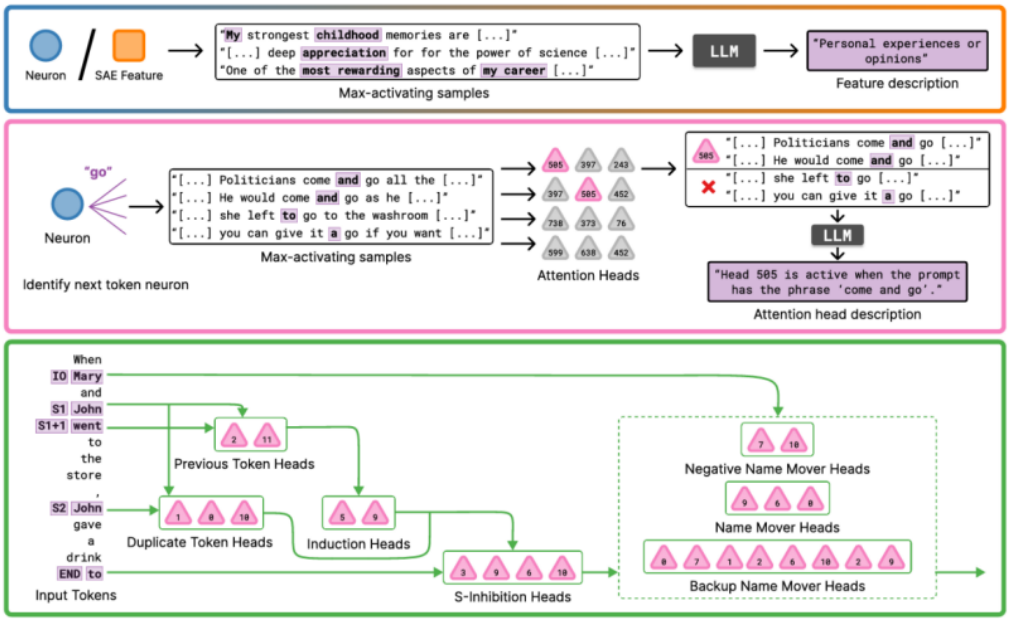


Figure 2: Overview of descriptions for model components (N neurons, A attention heads) and model abstractions (S SAE features, C circuits). The top panel shows a schematic example of an automatically generated feature description for a neuron or SAE feature, based on top-activating text samples (same process for both). The middle panel shows an example from Neo et al. (2024) of how attention head descriptions are generated: first a token-predicting neuron is identified, then prompts that highly activate it are found, the attention heads responsible for its activation are determined, and explanations for these heads are generated. The bottom panel shows a circuit from Wang et al. (2023a) implementing indirect object identification (IOI), where input tokens enter the residual stream and attention heads move information between streams, query/output arrows indicate where they write, key/value arrows where they read, and each class of head has an associated description.

主要亮点

- **首次全面梳理新兴领域：**是第一篇针对“通过概念描述解释语言模型”这一前沿领域的综述，系统性地整理了其核心方法、数据集和评估技术，为该领域的研究提供了结构化框架。

- **构建全面的方法与评估框架：**论文详细归纳了概念描述的生成方法（针对神经元、SAE 特征、电路等），并提出了一个包含预测模拟、因果干预、人工评估等五个维度的综合评估体系，为衡量解释质量提供了标准。

- **指明核心挑战与未来路线图：**论文深刻剖析了该领域面临的核心挑战（如神经元多义性、评估的因果性不足），并为未来研究提出了明确方向，包括将解释从组件扩展到电路、建立标准化基准等，旨在推动模型透明化的发展。

相关链接

Paper: https://arxiv.org/abs/2510.01048

● 论文六

Agentic Context Engineering: Evolving Contexts for Self-Improving Language Models

日期

2025-10-06

Agentic Context Engineering: Evolving Contexts for Self-Improving Language Models

Qizheng Zhang [1*] Changran Hu [2*] Shubhangi Upasani [2] Boyuan Ma [2] Fenglu Hong [2]
Vamsidhar Kamanuru [2] Jay Rainton [2] Chen Wu [2] Mengmeng Ji [2] Hanchen Li [3]
Urmish Thakker [2] James Zou [1] Kunle Olukotun [1]
[1] Stanford University [2] SambaNova Systems, Inc. [3] UC Berkeley [*] equal contribution
qizhengz@stanford.edu, changran.hu@sambanovasystems.com

内容提要

本文提出了一个名为 ACE（Agentic Context Engineering，即智能体化上下文工程）的框架，旨在解决当前大型语言模型（LLM）在上下文适应中所面临的关键挑战。现有方法常受困于“简洁性偏见”（为追求简练而牺牲领域洞察）和“上下文崩塌”（迭代重写导致细节随时间侵蚀）等问题。为应对此，ACE 将上下文视为一个持续演进的“剧本”（playbook），通过一个包含生成、反思和筛选的模块化流程，不断积累、提炼并组织策略。该框架采用结构化的增量更新机制以防止信息丢失，并能同时优化离线（如系统提示）与在线（如智能体记忆）的上下文。实验证明，ACE 在智能体和金融等领域特定基准测试中显著优于现有基线，并能在无需标记监督的情况下，仅利用执行反馈进行有效自适应，最终实现了一个可扩展、高效且能自我改进的 LLM 系统。

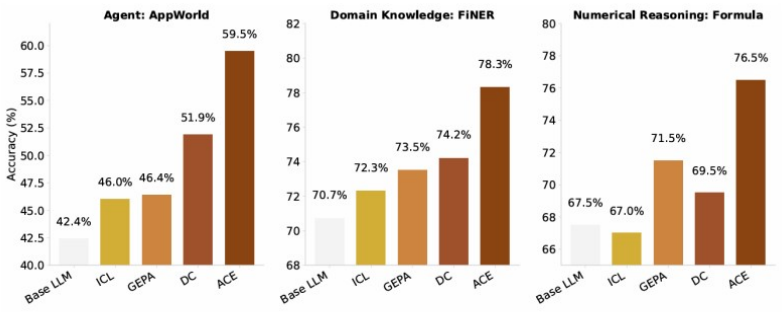


Figure 1: **Overall Performance Results.** Our proposed framework, ACE, consistently outperforms strong baselines across agent and domain-specific reasoning tasks.

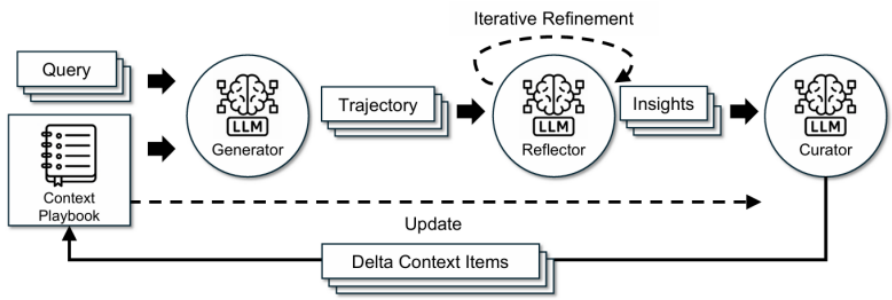


Figure 4: **The ACE Framework.** Inspired by Dynamic Cheatsheet, ACE adopts an agentic architecture with three specialized components: a Generator, a Reflector, and a Curator.

主要亮点

- **创新的“剧本”式上下文框架：**提出 Agentic Context Engineering (ACE) 框架，将上下文视为一个持续演进的“剧本”（playbook），通过不断积累、提炼和组织策略，有效解决了传统方法中的“简洁性偏见”和“上下文崩塌”问题。
- **模块化的增量更新机制：**采用“生成-反思-筛选”的模块化流程和结构化的增量更新，取代了低效且易导致信息丢失的整体重写方法，显著提升了上下文适应的效率、稳定性和可扩展性。
- **基于执行反馈的自我改进：**框架能够不依赖人工标注的正确答案，仅通过利用代码执行成功与否等自然反馈信号进行自我学习和优化，实现了真正的智能体自适应改进。
- **卓越的性能与效率：**在多个基准测试中，ACE 不仅大幅超越了现有方法（平均提升 8.6%-10.6%），还以更低的成本和延迟实现了上下文优化，并使得一个较小的开源模型在性能上媲美顶级的专有模型。

相关链接

Paper: https://arxiv.org/abs/2510.04618

● 论文七

Less is More: Recursive Reasoning with Tiny Networks

日期

2025-10-06

Less is More: Recursive Reasoning with Tiny Networks

Alexia Jolicoeur-Martineau
Samsung SAIL Montréal
alexia.j@samsung.com

内容提要

本文提出了 TRM（Tiny Recursive Model），一个用于解决困难推理任务的、基于微型网络的简化递归推理方法。解决如数独、迷宫和 ARC-AGI 等困难谜题任务对于大型语言模型（LLMs）一直是一个挑战，因为它们自回归的生成方式容易出错。虽然先前有 HRM（Hierarchical Reasoning Model）模型通过双网络分层递归取得了一定成功，但其结构复杂且依赖于不完全适用的理论。TRM 对此进行了简化和改进，仅使用一个单一的、层数极少（2 层）的微型网络。其核心思想是通过递归过程，迭代地优化一个“潜在答案”和一个“潜在推理”特征，从而逐步修正并逼近正确解。该方法摒弃了 HRM 复杂的双网络分层结构、生物学论证和对不动点定理的依赖，使得模型更简单、参数效率更高。实验结果表明，TRM 在参数量远小于（少于 0.01%）大型语言模型的情况下，在数独、迷宫和 ARC-AGI 等基准测试上取得了显著优于 HRM 乃至多数 LLM 的性能。该研究证明了“少即是多”的理念，为使用小型高效网络解决复杂推理问题提供了一条新的、有前景的技术路径。

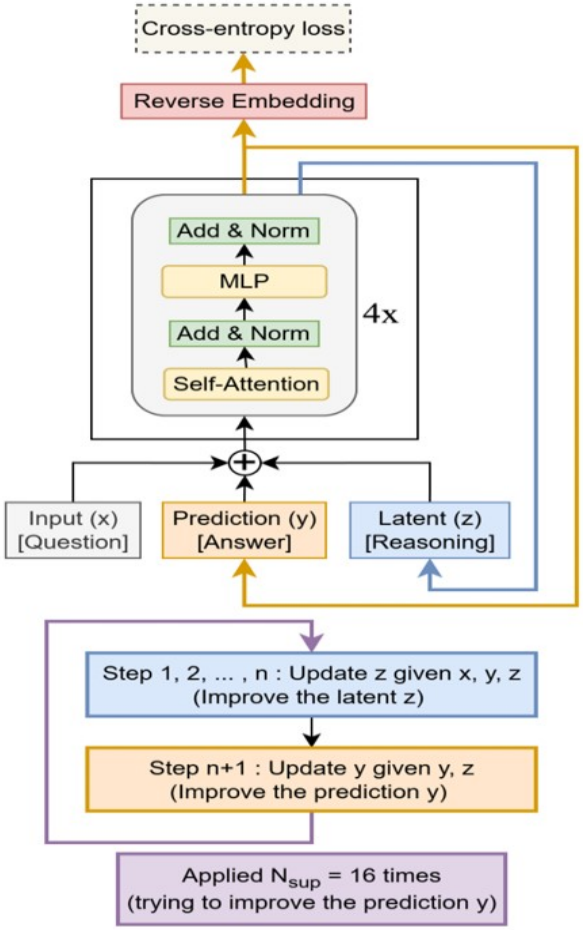


Figure 1. Tiny Recursion Model (TRM) recursively improves its predicted answer y with a tiny network. It starts with the embedded input question x and initial embedded answer y, and latent z. For up to $N_{sup} = 16$ improvements steps, it tries to improve its answer y. It does so by i) recursively updating n times its latent z given the question x, current answer y, and current latent z (recursive reasoning), and then ii) updating its answer y given the current answer y and current latent z. This recursive process allows the model to progressively improve its answer (potentially addressing any errors from its previous answer) in an extremely parameter-efficient manner while minimizing overfitting.

主要亮点

- **简化的递归推理框架：**提出了 TRM（Tiny Recursive Model），该模型使用单一微型网络，通过对“潜在答案”和“潜在推理”特征的递归迭代，逐步优化并生成最终解，摒弃了先前模型复杂的双网络分层结构。
- **极致的参数效率：**秉持“少即是多”的原则，TRM 仅用一个 2 层的微型网络（约 7M 参数）就在数独、迷宫和 ARC-AGI 等困难任务上取得了卓越性能，其参数量不到大型语言模型的 0.01%，却实现了更强的泛化能力。
- **精简的理论与训练：**该方法无需依赖复杂且不完全适用的不动点定理，也简化了训练中的自适应计算（ACT）机制，从而移除了额外的计算开销，使整个模型和训练过程更简单、高效且易于理解。

相关链接

Paper： https://arxiv.org/abs/2510.04871

● 论文八

Scientific Algorithm Discovery by Augmenting AlphaEvolve with Deep Research

日期

2025-10-07

SCIENTIFIC ALGORITHM DISCOVERY BY AUGMENTING ALPHAEVOLVE WITH DEEP RESEARCH

Gang Liu[1], **Yihan Zhu**[1] **Jie Chen**[2] **Meng Jiang**[1]
[1]University of Notre Dame [2] MIT-IBM Watson AI Lab, IBM Research
{gliu7, yzhu25, mjiang2}@nd.edu, chenjie@us.ibm.com

内容提要

本文提出了 DeepEvolve，一个通过深度研究来增强算法演化的科学发现智能体。现有的科学算法发现方法面临关键限制：单纯依赖大语言模型内部知识的算法演化（如 AlphaEvolve）在复杂领域很快会遇到性能瓶颈，而单纯的深度研究虽能提出想法却缺乏验证，导致方案不切实际或难以实现。DeepEvolve 将深度研究与算法演化相结合，在一个由反馈驱动的迭代循环中，整合了外部知识检索、跨文件代码编辑和系统性调试，从而将高质量的想法生成与可靠的执行结合起来。通过在化学、数学、生物学、材料和专利等领域的九个基准测试上的验证，该框架持续地改进了初始算法，生成了具有更高性能和效率的可执行新算法，为推进科学算法发现提供了一个可靠的框架。

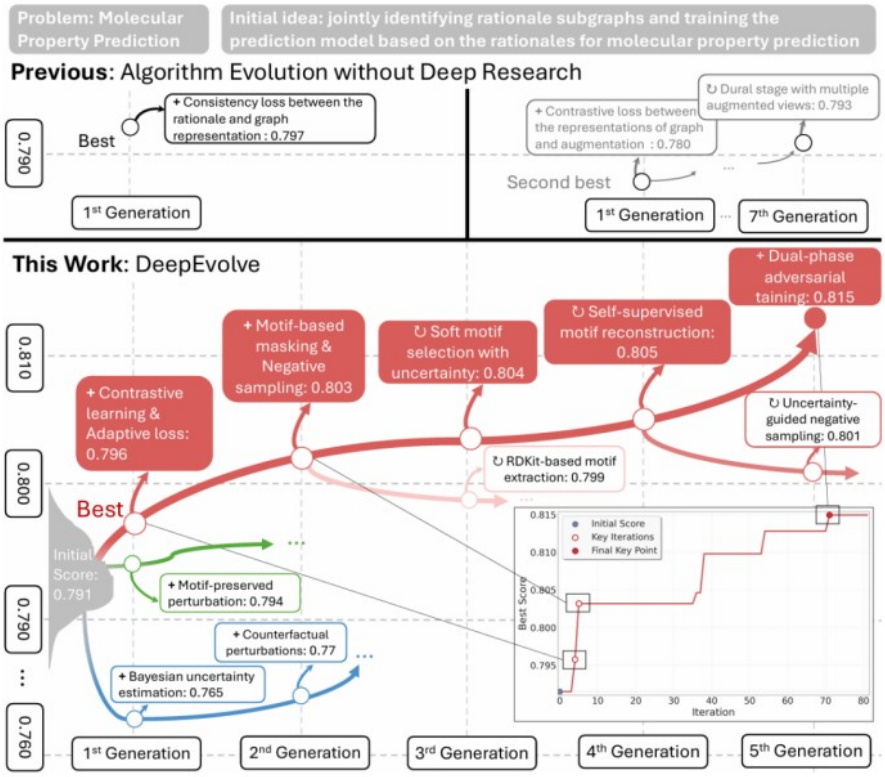


Figure 1: The top panel shows AlphaEvolve-style pure algorithm evolution without deep research, where the best improvement appears in the first generation and later iterations have marginal gains. The bottom panel shows DeepEvolve, which integrates deep research. DeepEvolve avoids shallow or excessively deep but unproductive evolutions, achieving sustained progress with clear performance jumps at key iterations. + denotes adding a new idea, and ↻ denotes refining a previous idea.

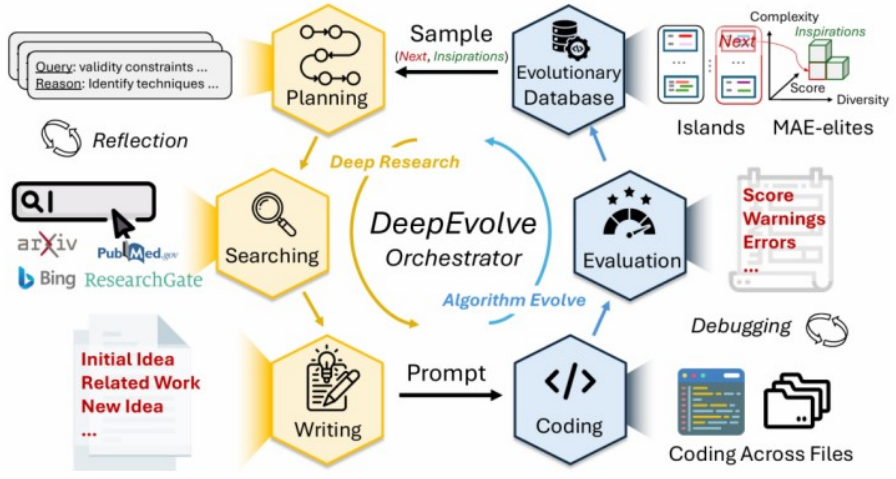


Figure 2: DeepEvolve is structured around six collaborative modules that alternate between deep research and algorithm evolution. Deep research generates informed hypotheses through planning, retrieval, and synthesis, while algorithm evolution translates these hypotheses into code, evaluates them, and applies evolutionary strategies for selection.

主要亮点

- **融合深度研究与算法演化**：提出 DeepEvolve 框架，将外部知识驱动的深度研究与算法演化相结合，解决了纯算法演化因知识局限而停滞，以及纯研究因缺乏验证而脱离实际的问题。
- **端到端的自动化算法实现**：整合了跨文件代码编辑、系统性调试和反馈驱动的迭代循环，能够将复杂的科学构想转化为可执行、可验证的代码，显著提高了算法实现的成功率。
- **跨领域的广泛验证与持续改进**：在化学、数学、生物学等九个不同科学领域的基准测试中，DeepEvolve 均能持续发现并生成性能优于初始版本的新算法，证明了其框架的通用性和有效性。

相关链接

Paper：https://arxiv.org/abs/2510.06056
GitHub：https://github.com/liugangcode/deepevolve

● 论文九

Webscale-RL: Automated Data Pipeline for Scaling RL Data to Pretraining Levels

日期

2025-10-07

Webscale-RL: Automated Data Pipeline for Scaling RL Data to Pretraining Levels

Zhepeng Cen[1,2], Haolin Chen[1], Shiyu Wang[1], Zuxin Liu[1], Zhiwei Liu[1], Ding Zhao[2], Silvio Savarese[1], Caiming Xiong[1], Huan Wang[1], Weiran Yao[1]

[1]Salesforce AI Research, [2]Carnegie Mellon University

内容提要

本文介绍了 Webscale-RL，一个用于将强化学习（RL）数据规模扩展至预训练水平的自动化数据管道。语言模型（LLM）的强化学习应用长期受限于数据瓶颈，即现有的 RL 数据集在规模和多样性上远不及网页级的预训练语料库，这阻碍了 RL 充分发挥其提升模型推理能力和解决训练-推理差距的潜力。为解决此问题，本文提出了 Webscale-RL 数据管道，该管道能够系统性地将大规模、多样化的预训练文档自动转换为数百万个可验证的、适用于 RL 训练的问答（QA）对。该流程包含数据筛选、领域分类与角色（persona）分配、可验证 QA 生成，以及质量与泄漏检查等多个阶段，旨在保证生成数据的质量、多样性和可靠性。基于该管道，作者构建了包含 120 万样本、覆盖 9 个以上领域的 Webscale-RL 数据集。实验证明，使用该数据集进行 RL 训练的模型性能显著优于在原始数据上进行持续预训练的基线模型，并且数据效率极高，能以少至 100 倍的令牌（tokens）达到同等性能，为将强化学习扩展至预训练规模并构建更强大、高效的语言模型提供了一条可行路径。

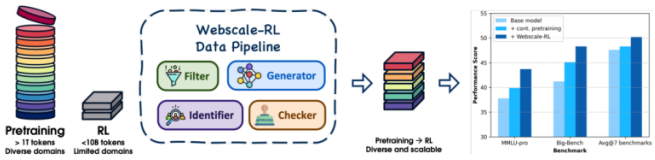


Figure 1: The scaling on LLM RL is fundamentally bottlenecked by the scarcity of high-quality RL data. While pretraining leverages >1T diverse web tokens, RL datasets remain limited to <10B tokens with limited diversity. We propose `Webscale-RL` data pipeline to fundamentally improve the scalability of RL data: we convert the pretraining corpora to verifiable query and ground-truth answer pairs, scaling RL data to pretraining levels while preserving the diversity. The experiments show that RL with `Webscale-RL` data is significantly more effective and efficient than continual pretraining and data refinement baselines.

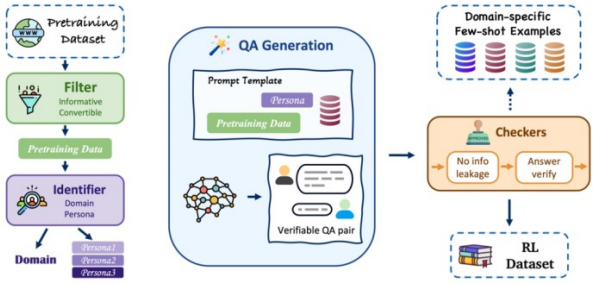


Figure 2: Overview of the `Webscale-RL` data pipeline that systematically converts large-scale pretraining data into RL data while preserving the scale and diversity of web data. The pipeline maintains a domain-specific demonstration library for few-shot examples for high quality generation and assigns multiple personas to each document to encourage reflecting different viewpoints. The generated QA pairs are verified for correctness and leakage prevention to ensure the reliability of the RL dataset.

主要亮点

- **自动化数据管道：**提出了 Webscale-RL 管道，一个可扩展的数据引擎，能将海量、多样化的预训练文档系统地转换为可供强化学习使用的、可验证的问答（QA）数据，解决了 RL 数据瓶颈问题。
- **大规模且多样化的 RL 数据集：**基于该管道构建了包含 120 万样本、覆盖超过 9 个领域的 Webscale-RL 数据集，其规模和多样性远超现有的 RL 数据集，为通用推理能力的训练提供了坚实基础。
- **显著的性能与数据效率提升：**实验证明，使用该数据集进行强化学习训练的模型性能显著优于持续预训练等基线方法，并能以高达 100 倍的数据效率实现同等性能，验证了该方法在提升模型能力和训练效率上的巨大潜力。

相关链接

Paper： https://arxiv.org/abs/2510.06499
GitHub: https://github.com/SalesforceAIResearch/PretrainRL-pipeline
Dataset: https://huggingface.co/datasets/Salesforce/Webscale-RL

● 论文十

SAM 3: Segment Anything with Concepts

日期

2025-10-12

SAM 3: SEGMENT ANYTHING WITH CONCEPTS

Anonymous authors
Paper under double-blind review

ABSTRACT

We present Segment Anything Model (SAM) 3, a unified model that detects, segments, and tracks objects in images and videos based on *concept prompts*, which we define as either short noun phrases (e.g., "yellow school bus"), image exemplars, or a combination of both. Promptable Concept Segmentation (PCS) takes such prompts and returns segmentation masks and unique identities for all matching object instances. To advance PCS, we build a scalable data engine that produces a high-quality dataset with 4M unique concept labels, including hard negatives, across images and videos. Our model consists of a vision backbone shared between an image-level detector and a memory-based video tracker. Recognition and localization are decoupled with a presence head, which significantly boosts detection accuracy. SAM 3 delivers a 2× gain over existing systems in both image and video PCS, and improves previous SAM capabilities in interactive visual segmentation tasks. We open source SAM 3 along with our new Segment Anything with Concepts (SA-Co) benchmark.

内容提要

本文提出了 SAM 3 (Segment Anything with Concepts)，一个旨在根据“概念提示”（如名词短语或图像范例）来检测、分割并跟踪图像与视频中所有匹配对象实例的统一模型。为实现这一目标，论文正式定义了“可提示概念分割”（Promptable Concept Segmentation, PCS）任务，它显著扩展了其前代模型仅能分割单个对象的能力。为了支持该任务，研究者构建了一个创新的、结合人类与 AI 标注员的大规模数据引擎，高效地生成了包含 400 万独特概念的高质量 SA-Co 数据集。模型架构上，SAM 3 采用检测器与跟踪器分离的设计，并引入一个新颖的“存在头部”（presence head）来解耦识别与定位，从而大幅提升了检测精度。同时，论文还发布了全新的 SA-Co 基准测试集，其规模远超现有基准，为评估 PCS 任务提供了坚实基础。实验结果显示，SAM 3 在图像和视频 PCS 任务上的性能均比现有系统提升了 2 倍，并在传统交互式分割任务上超越了 SAM 2。此外，SAM 3 还能与多模态大语言模型（MLLM）结合，形成“SAM 3 Agent”，以处理更复杂的推理分割请求。综上，SAM 3 通过定义新任务、构建高效数据引擎、提出创新模型架构并发布大规模基准，为可提示分割领域树立了新的里程碑。

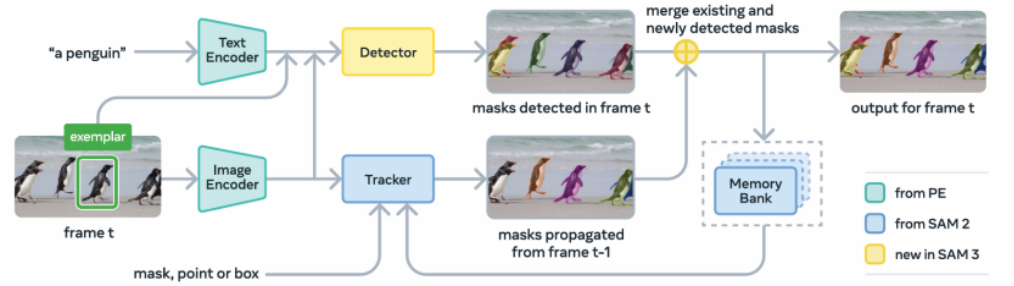


Figure 3: SAM 3 architecture overview. See Fig. 8 for a more detailed diagram.

主要亮点

- **从单个对象到概念分割的跨越**：SAM 3 定义了全新的“可提示概念分割”（PCS）任务，允许用户通过名词短语或图像范例作为提示，一次性分割出图像或视频中符合该概念的所有对象实例并进行跟踪，突破了以往模型一次仅能分割单个对象的局限。

- **创新的数据引擎与大规模基准**：论文构建了一个高效的人机协同数据引擎，创新地引入 AI 模型作为标注员和验证员，使标注吞吐量翻倍。基于此，不仅为模型训练生成了高质量数据集，还发布了全新的 SA-Co 基准，其概念数量比现有基准多 50 倍以上，为领域研究提供了宝贵资源。

- **解耦式模型架构与性能飞跃**：SAM 3 采用检测器与跟踪器分离的架构，并引入一个新颖的“存在头部”（presence head）将识别与定位任务解耦，显著提升了开放词汇检测的准确性。凭借此架构和高质量数据，SAM 3 在图像和视频概念分割任务上的性能相较于现有系统实现了翻倍。

相关链接

Paper：https://openreview.net/forum?id=r35clVtGzw

● 论文十一

Mamba-3: Improved Sequence Modeling using State Space Principles

日期

2025-10-12

MAMBA-3: IMPROVED SEQUENCE MODELING USING STATE SPACE PRINCIPLES

Anonymous authors
Paper under double-blind review

ABSTRACT

The recent scaling of test-time compute for LLMs has restricted the practical deployment of models to those with strong capabilities that can generate high-quality outputs in an inference-efficient manner. While current Transformer-based models are the standard, their quadratic compute and linear memory bottlenecks have spurred the development of sub-quadratic models with linear-scaling compute with constant memory requirements. However, many recent linear-style models lack certain capabilities or lag behind in quality, and even their linear-time inference is not hardware-efficient. Guided by an inference-first perspective, we introduce three core methodological improvements inspired by the state-space model viewpoint of linear models. We combine a: 1) more expressive recurrence, 2) complex state update rule that enables richer state tracking, and 3) multi-input, multi-output formulation together, resulting in a stronger model that better exploits hardware parallelism during decoding. Together with architectural refinements, our **Mamba-3** model achieves significant gains across retrieval, state-tracking, and downstream language modeling tasks. Our new architecture sets the Pareto-frontier for performance under a fixed inference budget and outperforms strong baselines in a head-to-head comparison.

内容提要

本文提出了 Mamba-3，一个基于状态空间模型（SSM）原理的改进序列建模架构。当前，序列建模领域面临着 Transformer 模型因其二次方计算复杂度和线性增长的内存（KV 缓存）瓶颈而带来的推理效率挑战，而现有的亚二次方模型（如 Mamba-2）虽然解决了效率问题，却常常在模型质量和特定能力（如状态追踪）上有所妥协。为此，Mamba-3 从推理效率优先的视角出发，引入了三项核心方法学改进：1）采用梯形离散化（Trapezoidal Discretization）构建了比以往方法更具表达能力的循环结构；2）引入复数值状态空间（Complex-valued SSMs），通过等效的数据依赖旋转位置编码（RoPE）增强了模型的状态追踪能力；3）提出了多输入多输出（MIMO）范式，通过将状态更新从外积转为矩阵乘法，显著提升了解码过程中的硬件计算效率。结合这些改进，Mamba-3 在保持线性时间复杂度和恒定内存需求的同时，在语言建模、检索和状态追踪等任务上取得了显著优于现有亚二次方模型的性能，为在固定推理预算下实现模型性能与效率的帕累托前沿（Pareto-frontier）树立了新的标杆。

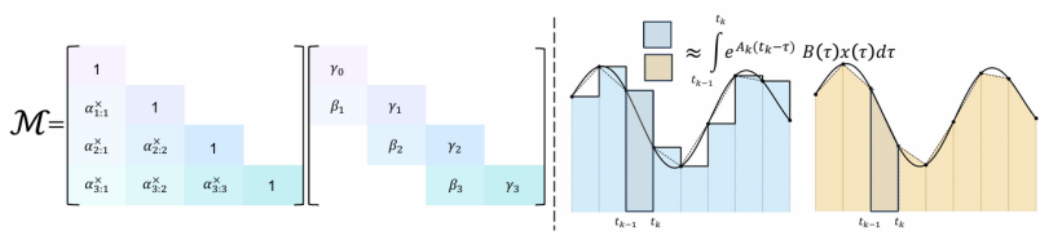


Figure 1: **Left:** The structured mask induced by the generalized trapezoid rule is a product of the decay and convolutional mask. **Right:** Euler (hold endpoint) vs trapezoidal rule (average endpoints).

主要亮点

- **梯形离散化改进：**采用比欧拉法更精确的梯形离散化方法，构建了表达能力更强的循环结构，在提升模型基础质量的同时，也使得在其他模型中常见的短卷积层变得不再必要。
- **复数值状态空间与状态追踪：**引入复数值状态空间，通过一种等效于数据依赖旋转位置编码（RoPE）的高效实现，使模型能够捕捉“旋转”动态，从而解决了 Mamba-2 等模型在奇偶校验、算术等状态追踪任务上的能力缺陷。
- **多输入多输出（MIMO）架构：**提出 MIMO 范式，将状态更新从外积升级为矩阵乘法，显著提高了自回归解码过程中的算术强度和硬件利用率，使得模型在不增加推理延迟和内存占用的情况下，获得了更强的性能。

相关链接

Paper： https://openreview.net/forum?id=HwCvaJOiCj

● 论文十二

BitNet Distillation

日期

2025-10-15

BitNet Distillation

Xun Wu Shaohan Huang Wenhui Wang Ting Song Li Dong Yan Xia Furu Wei†

Microsoft Research
https://aka.ms/GeneralAI

内容提要

本文介绍了 BitNet Distillation (BitDistill)，这是一个轻量级框架，旨在将现有的、全精度的大语言模型（LLMs）高效地微调为高度压缩的 1.58-bit（即三元权重 {-1, 0, 1}）模型，以用于特定的下游应用。这项工作解决了一个关键挑战：直接应用量化感知训练（QAT）将大模型微调至 1.58-bit 会导致显著的性能下降和差劲的可扩展性——即随着模型尺寸的增大，性能差距反而会扩大。BitDistill 通过一个三阶段的流程克服了这个问题：（1）使用 SubLN 模块进行架构优化以保证训练稳定性；（2）一个关键的持续预训练“热身”步骤，以使模型权重适应三元格式；（3）最后是一个基于蒸馏的微调阶段，该阶段利用 logits 蒸馏和多头注意力蒸馏来恢复全精度教师模型的性能。实验表明，BitDistill 使 1.58-bit 模型能够在性能上与其全精度版本相媲美，同时效率显著提高。

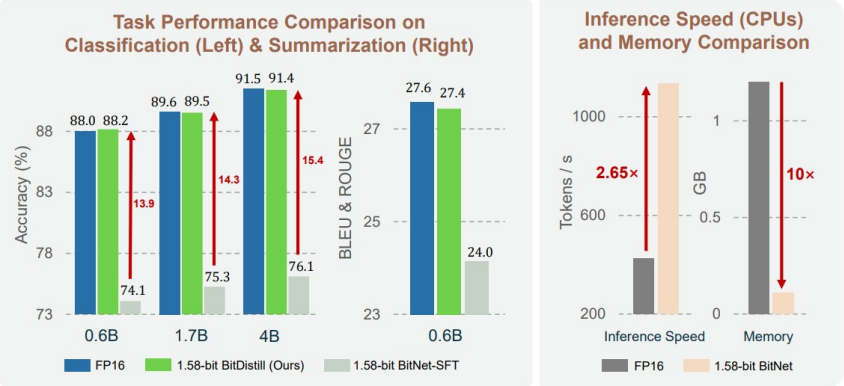


Figure 1: **Performance on downstream tasks across model size, with inference speed and memory efficiency comparison**. We observed that directly finetuning full-precision LLMs into 1.58-bit LLMs (denoted as 1.58-bit BitNet-SFT) leads to a notable performance gap compared to the FP16 baseline, and this gap remains or even widens as the model size increases. In contrast, `BitDistill` preserves scalability, resulting in performance comparable to full-precision counterparts across all model size, while reducing 10× memory usage and 2.65× faster inference on CPUs.

主要亮点

- **发现并解决关键的可扩展性问题：**该论文首次系统性地研究了为下游任务将预训练大模型微调至 1.58-bit 的问题，并发现了一个关键的可扩展性难题：随着模型尺寸的增加，全精度模型与 1.58-bit 模型之间的性能差距会越来越大。

- **新颖的三阶段蒸馏框架：**提出了一个名为 BitDistill 的定制化解决方案，它独特地结合了三项关键技术：（1）一个持续预训练的“热身”步骤，以有效缓解可扩展性问题；（2）多头注意力蒸馏（受 MiniLM 启发），以捕捉细粒度的结构知识；（3）使用 SubLN 模块来确保稳定的优化过程。

- **以极致效率实现 FP16 级别性能：**所提出的方法使 1.58-bit 模型在各种下游任务上，能达到与其全精度（FP16）版本相媲美的性能。与此同时，它还带来了巨大的效率提升，包括高达 10 倍的内存节省和在 CPU 上 2.65 倍的推理加速，使其非常适合在资源受限的设备上部署。

相关链接

Paper： https://arxiv.org/abs/2510.13998
GitHub： https://github.com/microsoft/BitNet

● 论文十三

Reasoning with Sampling: Your Base Model is Smarter Than You Think

日期

2025-10-16

Reasoning with Sampling: Your Base Model is Smarter Than You Think

Aayush Karan[1], Yilun Du[1]

[1]Harvard University

Website Code

内容提要

本文提出了一种无需训练的采样算法，旨在挑战强化学习（RL）在提升大语言模型（LLM）推理能力方面的主导地位。研究的核心问题是：强大的推理能力是否只能通过 RL 后期训练获得，还是说基础模型本身就具备、只是需要更优的推理时（inference-time）方法来激发？作者提出了“幂采样”（Power Sampling）算法，该算法受马尔可夫链蒙特卡洛（MCMC）方法的启发，通过迭代式地重新采样来近似地从基础模型的“幂分布”中抽取样本。实验证明，该方法能在多种推理任务（如 MATH500, HumanEval, GPQA）上显著提升基础模型的单次推理性能，其效果媲美甚至超越了经过 RL 训练的模型，并且无需任何额外训练、精选数据集或验证器。

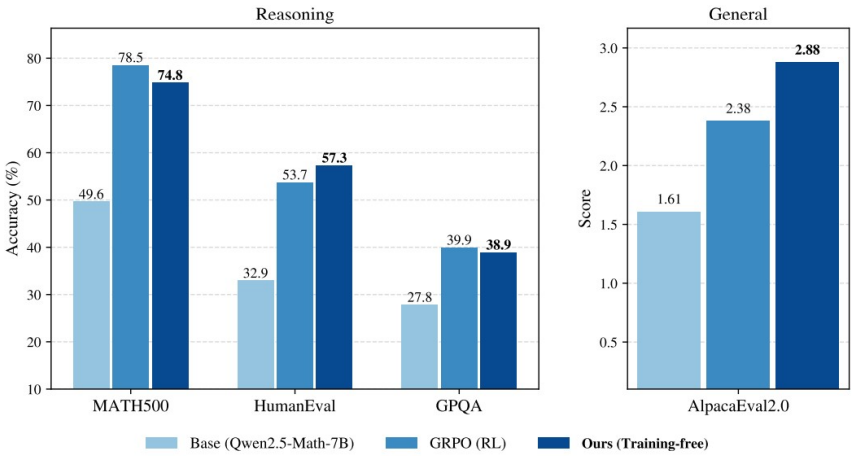


Figure 1: **Our sampling algorithm can match and outperform RL-posttraining.** Left: we compare our sampling algorithm (ours) against the base model (base) and RL-posttraining (GRPO) on three *verifiable reasoning* tasks (MATH500, HumanEval, GPQA). Right: we compare them on an *unverifiable general task* (AlpacaEval2.0). Our algorithm achieves comparable performance to GRPO within the posttraining domain (MATH500) but can *outperform* on out-of-domain tasks such as HumanEval and AlpacaEval.

主要亮点

- **无需训练即可媲美强化学习**：一个纯粹在推理阶段通过巧妙采样运行的算法，可以在单次推理任务上达到与经过复杂且昂贵的强化学习（GRPO）后期训练相当甚至更好的性能。例如，在 HumanEval 任务上，该方法将 Phi-3.5-mini-instruct 模型的准确率从 21.3% 提升至 73.2%，远超 RL 方法的 13.4%。

- **保留样本多样性，解决 RL 痛点**：与强化学习方法普遍存在“模式崩溃”、导致生成样本多样性下降的问题不同，幂采样算法成功地保留了基础模型的多样性。在 pass@k（多样本通过率）评估中，当 k>1 时，该算法的性能远超 RL 模型，并能达到基础模型的上限，实现了单次推理高准确率与多次推理高多样性的“两全其美”。

- **基于“幂分布”的创新采样理论**：该方法在理论上将推理问题与从“幂分布”（p^α）采样联系起来，并清晰地论证了它与传统低温（low-temperature）采样的区别。幂分布采样能更好地“规划”未来，倾向于选择那些后续路径虽少但可能性高的词元（token），从而有效避开导致推理失败的“关键陷阱”（pivotal tokens）。

相关链接

Paper：https://arxiv.org/abs/2510.14901
GitHub: https://github.com/aakaran/reasoning-with-sampling

● 论文十四

From Pixels to Words -- Towards Native Vision-Language Primitives at Scale

日期

2025-10-16

FROM PIXELS TO WORDS – TOWARDS NATIVE VISION-LANGUAGE PRIMITIVES AT SCALE

Haiwen Diao[1] **Mingxuan Li**[2] **Silei Wu**[3] **Linjun Dai**[3]
Xiaohua Wang[2] **Hanming Deng**[3] **Lewei Lu**[3] **Dahua Lin**[3] **Ziwei Liu**[1]
[1]S-Lab, Nanyang Technological University
[2]Xi'an Jiaotong University [3]SenseTime Research

内容提要

本文介绍了 NEO，一个全新的原生视觉语言模型（Native VLM）家族。该模型旨在从根本上解决传统模块化 VLM（拼接独立视觉与语言模型）所带来的结构复杂、对齐成本高昂等问题，并克服现有原生 VLM 在效率和稳定性上的瓶颈。NEO 通过引入统一的“原生 VLM 基元”（Native VLM Primitives）来构建单一的、整体化的模型架构，其核心创新包括“原生多头注意力机制”（MHNA）和“原生旋转位置编码”（Native-RoPE），以实现像素与文本在共享语义空间内的深度融合与协同推理。通过创新的“预缓冲层-后语言模型”分阶段训练策略，NEO 仅使用 390M 图文数据便能从零开始高效学习视觉感知，其综合性能在多个基准测试中成功媲美了同等规模的顶尖模块化 VLM。

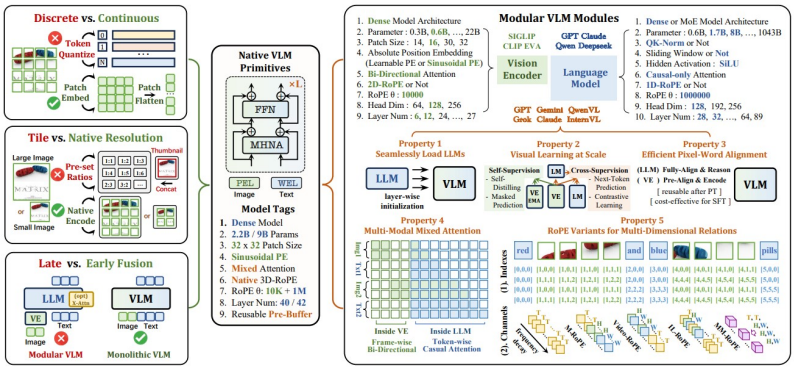


Figure 1: Overview of our native vision-language frameworks, which project arbitrary-resolution images into a continuous latent space, integrating the virtues of modular VLM architectures and enabling efficient vision-language encoding, alignment, and interaction in an early-fusion manner.

主要亮点

- **原生统一架构范式**：NEO 摒弃了将预训练视觉编码器（VE）与大型语言模型（LLM）连接的传统模块化思路，提出了一种从第一性原理构建的原生 VLM 架构。其核心是统一的“原生 VLM 基元”，该基元在单一模块内无缝集成了视觉-语言的编码、对齐与推理功能，为构建真正一体化的多模态大模型提供了全新的、更简洁的路径。

- **创新的原生基元设计**：NEO 引入了多项关键技术以增强跨模态交互。原生多头注意力机制（MHNA）巧妙地在同一模型层中对图像使用双向注意力、对文本使用因果注意力，从而高效建模复杂的跨模态依赖关系。原生旋转位置编码（Native-RoPE）通过解耦时间（T）、高度（H）和宽度（W）维度并为其分配不同频率，在不破坏预训练 LLM 语言能力的同时，实现了对空间关系的精细感知。

- **高效的训练策略与卓越性能**：NEO 采用创新的“预缓冲层（Pre-Buffer）-后语言模型（Post-LLM）”分阶段训练策略，使模型能够从零开始高效学习视觉知识，同时完整地吸收并保留 LLM 强大的推理能力。凭借此设计，NEO 仅用少量数据（390M 图文对）就在多个主流基准测试上取得了与使用数十亿数据训练的顶尖模块化 VLM 相媲美的性能，充分证明了其架构的优越性、可扩展性和数据效率。

相关链接

Paper： https://arxiv.org/abs/2510.14979
GitHub: https://github.com/EvolvingLMMs-Lab/NEO

● 论文十五

BLIP3o-NEXT: Next Frontier of Native Image Generation

日期

2025-10-17

BLIP3o-NEXT: Next Frontier of Native Image Generation

Jiuhai Chen[1,2*] Le Xue[1*] Zhiyang Xu[3*] Xichen Pan[4] Shusheng Yang[4]
Can Qin[1] An Yan[1] Honglu Zhou[1] Zeyuan Chen[1] Lifu Huang[6] Tianyi Zhou[2]
Junnan Li[1] Silvio Savarese[1‡] Caiming Xiong[1‡] Ran Xu[1‡]

[1]Salesforce Research
[2]University of Maryland [3]Virginia Tech [4]New York University [6] UC Davis

[*]Equal Contribution. [‡]Corresponding Authors.

内容提要

本文介绍了 BLIP3o-NEXT，一个在 BLIP3 系列中完全开源的基础模型，旨在推动原生图像生成技术的新前沿。BLIP3o-NEXT 在单一架构内统一了文生图（text-to-image）和图像编辑两大功能，并展现出强大的性能。该模型采用了一种“自回归 + 扩散”（Autoregressive + Diffusion）的混合架构：首先，一个自回归模型根据多模态输入（文本和参考图像）生成离散的图像标记（image tokens）；然后，这些标记的隐藏状态（hidden states）被用作一个扩散模型的条件信号，以生成高保真度的最终图像。这种设计结合了自回归模型的推理、指令遵循能力与扩散模型的精细细节渲染能力，实现了语义连贯性与照片级真实感的新高度。

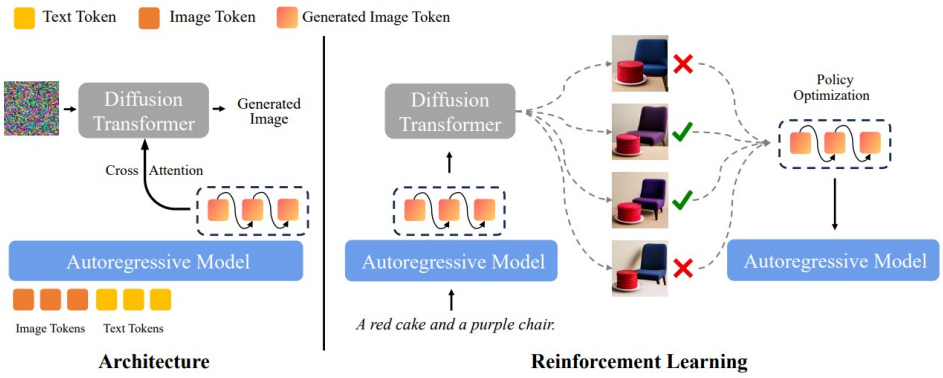


Figure 1: The architecture of BLIP3o-NEXT (left) and its reinforcement learning pipeline (right). BLIP3o-NEXT adopts an Autoregressive (AR) + Diffusion design, where the AR module autoregressively generates image conditions for the diffusion model. The model is jointly optimized with both AR and diffusion objectives. During reinforcement learning, rollouts are rendered from the diffusion transformer, and policy optimization is performed directly on the AR model, enabling seamless integration with existing RL infrastructures originally developed for language models.

主要亮点

- **创新的“自回归+扩散”架构与高效强化学习：**模型采用 AR+Diffusion 架构，并创新性地将强化学习（RL）直接应用于自回归模块。这种方法能够无缝集成现有为语言模型开发的 RL 框架，高效地提升了模型的文本渲染和复杂指令遵循能力。

- **统一的图像生成与编辑能力，并系统性提升编辑一致性：**BLIP3o-NEXT 在单一模型中实现了图像生成和编辑。为解决编辑任务中保持一致性的核心挑战，论文系统地研究并采用了多种策略，包括引入图像重建任务，以及将参考图像的 VAE 特征注入到扩散过程中，显著增强了编辑后图像的保真度。

- **完全开源，推动社区发展：**秉承 BLIP3 系列的开源理念，作者完全开放了 BLIP3o-NEXT 的所有资源，包括预训练和后训练的模型权重、数据集、详细的训练和推理代码以及评估流程，确保了研究的可复现性，旨在促进原生图像生成领域的持续进步。

相关链接

Paper： https://arxiv.org/abs/2510.15857
GitHub： https://github.com/JiuhaiChen/BLIP3o

● 论文十六

DeepAnalyze: Agentic Large Language Models for Autonomous Data Science

日期

2025-10-19

DeepAnalyze: Agentic Large Language Models for Autonomous Data Science

Shaolei Zhang [1] Ju Fan [1] Meihao Fan [1] Guoliang Li [2] Xiaoyong Du [1]

内容提要

本文介绍了 DeepAnalyze-8B，这是首个为实现“自主数据科学”而设计的智能体大语言模型（agentic LLM）。该模型能够端到端地自动化整个数据科学流程，从原始数据源出发，最终生成分析师级别的深度研究报告。为了克服传统基于工作流的数据智能体在处理高度复杂任务时的局限性，作者提出了一种“基于课程的智能体训练范式”（curriculum-based agentic training paradigm）。该范式模仿人类数据科学家的学习路径，让模型在真实环境中逐步掌握并融合多种数据科学能力。同时，为了解决高质量训练数据稀缺的问题，研究还提出了一个“数据驱动的轨迹合成框架”（data-grounded trajectory synthesis framework）。实验证明，仅有 80 亿参数的 DeepAnalyze 在性能上超越了基于最先进专有大模型构建的工作流智能体。

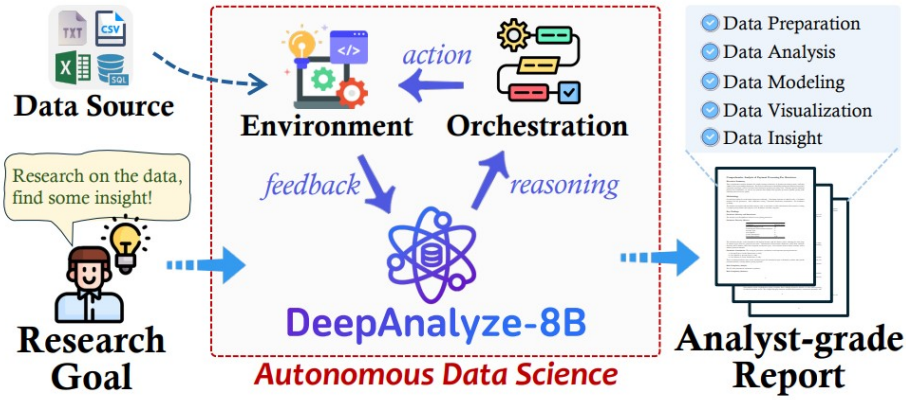


Figure 1. DeepAnalyze-8B is the first end-to-end agentic LLM that achieves autonomous data science, supporting entire data science pipeline and open-ended data research.

主要亮点

- **首个为自主数据科学设计的智能体大模型：**DeepAnalyze-8B 是第一个能够端到端完成从数据准备、分析、建模、可视化到生成深度研究报告全流程的智能体模型。它不再依赖预定义的工作流，而是通过自主规划和与环境的交互来解决开放式的数据研究问题。

- **创新的“基于课程的智能体训练”范式：**为解决数据科学任务的高复杂性导致的奖励稀疏问题，该研究模仿人类“从易到难”的学习过程，设计了“先单项能力微调，再多项能力智能体训练”的两阶段课程。这使得模型能逐步掌握并整合推理、代码生成、结构化数据理解等多种复杂技能。

- **数据驱动的轨迹合成框架，解决训练数据稀缺问题：**为了给智能体训练提供高质量的“示范”，论文提出了一个多智能体协作框架，能自动从现有的结构化数据源（如数据库）中合成高质量的推理和交互轨迹。这有效解决了数据科学领域缺乏长链、多步问题解决过程数据的难题。

相关链接

Paper： https://arxiv.org/abs/2510.16872
GitHub： https://github.com/ruc-datalab/DeepAnalyze
Model： https://huggingface.co/RUC-DataLab/DeepAnalyze-8B

● 论文十七

GS-World: An Efficient, Engine-driven Learning Paradigm for Pursuing Embodied Intelligence using World Models of Generative Simulation

日期

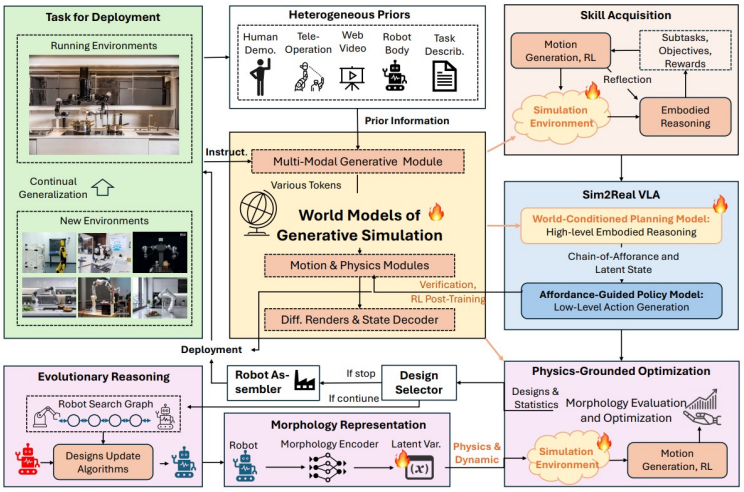
2025-10-19

GS-World: An Engine-driven Learning Paradigm for Pursuing Embodied Intelligence using World Models of Generative Simulation

Guiliang Liu, Yueci Deng, Zhen Liu, Kui Jia*
School of Data Science, The Chinese University of Hong Kong, Shenzhen
liuguiliang@cuhk.edu.cn, yuecideng@link.cuhk.edu.cn,
{zhenliu,kuijia}@cuhk.edu.cn

内容提要

本文提出了一种名为 GS-World（生成式模拟世界模型）的新型引擎，并基于此构建了一个用于追求具身智能的高效学习范式。该范式旨在解决当前具身智能研究中因缺乏大规模、多模态、物理相关的训练数据而导致的“效率定律”瓶颈。GS-World 通过生成式学习来构建一个完全遵循物理规律的模拟世界，包括 3D 资产、环境以及支配它们动态交互的物理规则。基于这个引擎，论文提出了一种名为“引擎驱动的 Sim2Real VLA”的自动化学习流程，该流程涵盖了数据生成、视觉-语言-动作（VLA）模型训练、验证和部署的全过程。此范式不仅能够高效地训练出具身智能体，还支持对智能体形态进行任务导向的优化。



主要亮点

- **提出“效率定律”并构建 GS-World 引擎：**论文首先分析了具身智能领域难以应用“规模定律”的原因，并提出了一个更关键的“效率定律”，即模型性能与数据生成效率直接相关。为解决此问题，论文提出了 GS-World（生成式模拟世界模型）引擎。该引擎能够生成一个完全符合物理规律的虚拟世界，为具身智能的学习提供了无限的、高质量的、物理精确的训练数据来源，从根本上解决了数据稀缺的难题。

- **创新的“引擎驱动的 Sim2Real VLA”学习范式：**基于 GS-World 引擎，论文构建了一个自动化、闭环的学习流程。该流程实现了从数据生成、VLA 模型训练、在线验证到真实世界部署的无缝衔接。这种“引擎驱动”的模式取代了传统依赖于手动收集数据的“数据驱动”模式，极大地提升了学习效率和可扩展性，使得持续、自主地提升具身智能成为可能。

- **物理精确性与形态和策略的协同进化：**与仅追求视觉真实感的视频生成模型不同，GS-World 强调物理精确性，能够模拟真实的力学、接触和动态交互，从而保证了学习到的策略在真实世界中的有效性（Sim2Real）。更进一步，由于整个学习流程是可微的，该范式不仅能优化控制策略（“大脑”），还能反向优化机器人的形态结构（“身体”），实现了形态与智能的协同进化，从而能够为特定任务自主设计出最优的机器人形态。

相关链接

Paper： https://openreview.net/pdf?id=E23dMeRpE0

● 论文十八

DeepSeek-OCR: Contexts Optical Compression

日期

2025-10-21

DeepSeek-OCR: Contexts Optical Compression

Haoran Wei, Yaofeng Sun, Yukun Li

DeepSeek-AI

内容提要

本文介绍了 DeepSeek-OCR，一项旨在探索通过光学二维映射（optical 2D mapping）来压缩长文本上下文可行性的初步研究。该模型由一个创新的视觉编码器 DeepEncoder 和一个 DeepSeek3B-MoE 解码器构成。DeepEncoder 作为核心引擎，被设计用于在高分辨率输入下保持低激活内存和高压缩比，以生成数量可控的视觉令牌。实验结果表明，当压缩比低于 10 倍时，模型可以实现 97% 的 OCR 解码精度；即使在 20 倍的高压缩比下，精度仍能保持在 60% 左右。这不仅为大语言模型的长上下文处理和记忆遗忘机制等领域提供了新思路，同时 DeepSeek-OCR 在实际应用中也展现了极高价值，在 OmniDocBench 基准上使用远少于同类模型的视觉令牌便取得了领先性能。

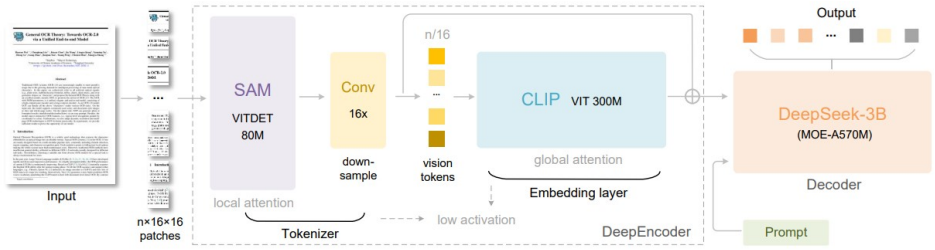


Figure 3 | The architecture of DeepSeek-OCR. DeepSeek-OCR consists of a DeepEncoder and a DeepSeek-3B-MoE decoder. DeepEncoder is the core of DeepSeek-OCR, comprising three components: a SAM [17] for perception dominated by window attention, a CLIP [29] for knowledge with dense global attention, and a 16× token compressor that bridges between them.

主要亮点

- **上下文光学压缩的可行性验证：**DeepSeek-OCR 首次对“视觉-文本”令牌的压缩比进行了全面的量化分析，将长文本处理问题重新定义为一个压缩-解压缩任务。在低于 10 倍的压缩比下实现了 97% 的 OCR 解码精度，在 20 倍压缩比下仍保持约 60% 的精度，为解决大语言模型长上下文难题提供了全新的光学压缩思路。
- **创新的 DeepEncoder 架构：**论文提出了一种新颖的视觉编码器 DeepEncoder，它通过一个 16 倍卷积压缩器，将以窗口注意力为主的 SAM 模块和以全局注意力为主的 CLIP 模块串联起来。这种设计能够在处理高分辨率图像时，有效控制激活内存和输出的视觉令牌数量，解决了现有视觉编码器在处理密集信息图像时的普遍痛点。
- **高效率下实现顶尖性能：**在 OmniDocBench 基准上，DeepSeek-OCR 仅用 100 个视觉令牌就超越了使用 256 个令牌的 GOT-OCR2.0；并用少于 800 个令牌击败了需要 6000+ 令牌的 MinerU2.0。此外，该模型具备强大的生产能力，单张 A100-40G GPU 每天可生成超过 20 万页的训练数据，展示了其巨大的研究潜力和实际应用价值。

相关链接

Paper：https://arxiv.org/abs/2510.18234
GitHub：https://github.com/deepseek-ai/DeepSeek-OCR
Model：https://huggingface.co/deepseek-ai/DeepSeek-OCR

● 论文十九

Pico-Banana-400K: A Large-Scale Dataset for Text-Guided Image Editing

日期

2025-10-22

Pico-Banana-400K: A Large-Scale Dataset for Text-Guided Image Editing

Yusu Qian, Eli Bocek-Rivele, Liangchen Song, Jialing Tong, Yinfei Yang, Jiasen Lu*, Wenze Hu*, Zhe Gan*

Apple

*Senior authors

内容提要

本文介绍了 Pico-Banana-400K，一个包含约 40 万个样本的大规模、高质量、开放的文本指导图像编辑数据集。为解决现有数据集在真实性、质量和多样性上的局限，该研究利用先进的 Nano-Banana 模型对 OpenImages 数据集中的真实照片进行编辑，并采用 Gemini-2.5-Pro 模型作为自动评估员进行严格的质量控制。Pico-Banana-400K 的独特之处在于其系统化的构建方法，它不仅涵盖了 35 种细粒度的编辑类型，还包含了三个为复杂研究场景设计的专业子集：用于序列化编辑研究的多轮对话编辑数据、用于对齐和奖励模型训练的偏好数据，以及用于指令改写的长短指令对。该数据集旨在为下一代图像编辑模型的训练和基准测试提供一个坚实的基础。

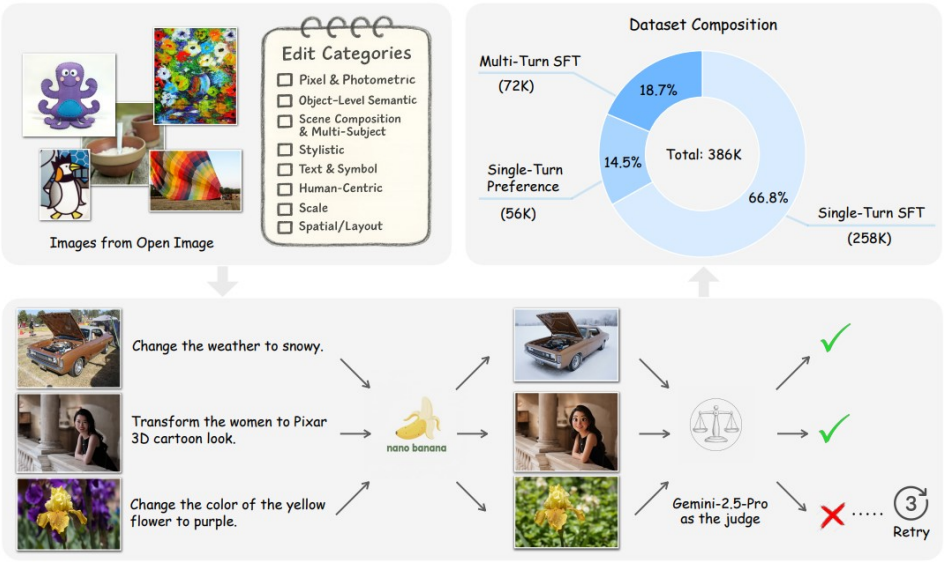


Figure 1 Pico-Banana-400K dataset overview. The pipeline (bottom) shows how diverse OpenImages inputs are edited using Nano-Banana and quality-filtered by Gemini-2.5-Pro, with failed attempts automatically retried. The dataset comprises 386K examples across single-turn SFT (66.8%), preference pairs (14.5%), and multi-turn sequences (18.7%), organized by our comprehensive edit taxonomy (top left).

主要亮点

- **大规模高质量真实图像数据集：**基于真实世界图像，利用顶尖模型（Nano-Banana 生成，Gemini-2.5-Pro 质检）构建了一个包含约 40 万样本的高质量、多样化（35 种编辑类型）的公开数据集。
- **为前沿研究设计的丰富数据子集：**创新性地提供了三个专业子集，分别用于研究序列化编辑（多轮对话数据）、模型对齐（偏好数据）和指令理解（长短指令对）。
- **创新的自动化数据构建流程：**采用“一个模型生成，一个模型质检”的自动化闭环流程，不仅保证了数据质量和可扩展性，还将失败样本转化为有价值的偏好数据，为社区提供了高效的数据生产范式。

相关链接

Paper: https://arxiv.org/abs/2510.19808

Github： https://github.com/apple/pico-banana-400k

● 论文二十

olmOCR 2: Unit Test Rewards for Document OCR

日期

2025-10-22

内容提要

本文介绍了 olmOCR 2，一个用于将数字化印刷文档 (如 PDF) 转换为整洁、自然顺序纯文本的顶尖 OCR 系统。olmOCR 2 的核心是一个名为 olmOCR-2-7B-1025 的 7B 视觉语言模型（VLM），其创新之处在于使用了“带可验证奖励的强化学习”（RLVR）进行训练。为了规模化地实现这一点，作者开发了一个合成数据流程，该流程能将真实文档渲染为带有已知真值（ground-truth）的 HTML，并从中自动生成可验证的二元单元测试用例。通过在这些单元测试上进行强化学习训练（GRPO），olmOCR 2 在 olmOCR-Bench 基准上取得了当前最佳性能，尤其在数学公式转换、表格解析和多栏布局等方面的改进最为显著。

	olmOCR-Bench score	Release date	Model weights	Training data	Training code	Inference code	Model license
OpenAI GPT-4o	68.9 ± 1.1	May 2024	✗	✗	✗	✗	✗ [6]
Qwen 2 VL 7B	31.5 ± 0.9	Aug 2024	✓	✗	✗	✓	✓ [1]
Gemini Flash 2	57.8 ± 1.1	Dec 2024	✗	✗	✗	✗	✗ [6]
Qwen 2.5 VL 7B	65.5 ± 1.2	Feb 2025	✓	✗	✗	✓	✓ [1]
Mistral OCR API	72.0 ± 1.1	Mar 2025	✗	✗	✗	✗	✗ [6]
MinerU `1.3.10`	61.5 ± 1.1	Apr 2025	✓	✗	✗	✓	! [4]
Nanonets OCR S	64.5 ± 1.1	Jun 2025	✓	✗	✗	✓	? [5]
MonkeyOCR Pro 3B	75.8 ± 1.0*	Jun 2025	✓	✗	✗	✓	? [5]
Infinity-Parser 7B	79.1 ± ?*	Jun 2025	✓	✓	✗	✓	✓ [1]
dots.OCR	79.1 ± 1.0*	Jul 2025	✓	✗	✓	✓	✓ [2]
Marker `1.10.1`	76.1 ± 1.1	Sep 2025	✓	✗	✗	✓	! [3]
MinerU `2.5.4`	75.2 ± 1.1*	Sep 2025	✓	✗	✗	✓	! [4]
PaddleOCR-VL	80.0 ± 1.0*	Oct 2025	✓	✗	✓	✓	✓ [1]
Nanonets OCR2 3B	69.5 ± 1.1	Oct 2025	✓	✗	✗	✓	? [5]
DeepSeek-OCR	75.7 ± 1.0	Oct 2025	✓	✗	✗	✓	✓ [2]
Infinity-Parser 7B	**82.5 ± ?***	Oct 2025	✓	✗	✗	✓	✓ [1]
Chandra OCR `0.1.0`	**83.1 ± 0.9***	Oct 2025	✓	✗	✗	✓	! [3]
OLMOCR	68.2 ± 1.1	Feb 2025	✓	✓	✓	✓	✓ [1]
OLMOCR 2	**82.4 ± 1.1**	Oct 2025	✓	✓	✓	✓	✓ [1]

Table 1 Comparison of OLMOCR 2 and other OCR systems. OLMOCR 2 achieves state-of-the-art performance while maintaining fully open data, model, and code. Open-source licenses: [1]Apache 2.0, [2]MIT; open licenses with usage restrictions: [3]OpenRAIL-M, [4]AGPL v3; [5]license not specified; [6]API access only after accepting ToS. Results are fully reproduced by ourselves, except those marked with * which are reported by their authors.

主要亮点

- **以单元测试作为强化学习奖励**：创新性地使用二元（通过/失败）单元测试作为奖励信号进行 RLVR 训练，替代了传统的编辑距离。这种方法能更准确地评估 OCR 输出的语义正确性，尤其在处理浮动元素、数学公式和阅读顺序等复杂情况时。
- **可扩展的合成数据与测试生成流程**：开发了一个自动化流程，能将真实世界文档（如 PDF）渲染为带有已知真值的 HTML，并从中自动提取大量的单元测试用例。这成功解决了为 RL 训练大规模创建高质量、带标签数据的难题。
- **顶尖性能与完全开放**：olmOCR 2 在 olmOCR-Bench 基准上达到了顶尖水平，相比初版有 14.2 分的显著提升。同时，项目坚 Demo 持完全开放的原则，其模型、代码和数据均在宽松的开源许可下发布，有力地推动了社区的研究。

相关链接

Paper: https://arxiv.org/abs/2510.19817

Github： https://github.com/allenai/olmocr

Demo: https://olmocr.allenai.org/

Model： https://huggingface.co/allenai/olmOCR-2-7B-1025

● 论文二十一

Discovering state-of-the-art reinforcement learning algorithms

日期

2025-10-22（Nature 录用论文）

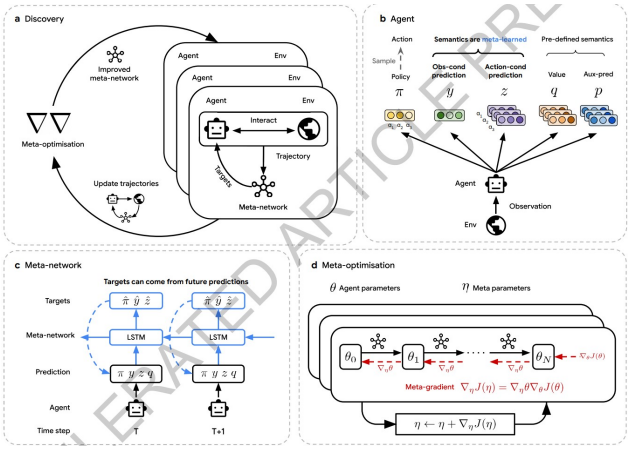
Discovering state-of-the-art reinforcement learning algorithms

Received: 11 December 2024
Accepted: 15 October 2025

Junhyuk Oh, Greg Farquhar, Iurii Kemaev, Dan A. Calian, Matteo Hessel, Luisa Zintgraf, Satinder Singh, Hado van Hasselt & David Silver

内容提要

本文介绍了一种能够自动发现全新、高性能强化学习（RL）算法的方法。该研究旨在解决一个长期存在的挑战：与动物通过进化发现学习机制不同，人工智能体通常依赖于人类手工设计的学习规则。作者提出的方法通过“元学习”（meta-learning），让一个大型智能体种群在大量复杂多样的环境中进行交互，并从它们的集体经验中学习出一个 RL 更新规则。该方法的核心是一个作为 RL 规则的“元网络”（meta-network）。它不预设价值函数或策略梯度等概念，而是学习为智能体的策略和一组可被发现的、无预定语义的预测向量生成更新目标。元网络的参数通过元梯度进行优化，以最大化智能体种群的累积奖励。最终被发现的算法，命名为 DiscoRL，不仅自主地重新发现了像“自举”（bootstrapping）这样的基本 RL 概念，还发展出了独特的、新的预测语义。



主要亮点

- **从第一性原理发现超越人类设计的顶尖 RL 算法：**这项工作标志着一个重要的里程碑，它首次通过自动发现的方式，创造出一个在性能上达到顶尖水平的 RL 算法。被发现的 DiscoRL 在公认的 Atari 基准测试中，超越了包括 MuZero 在内的所有现有手动设计的算法，证明了机器发现的算法能够与甚至优于人类设计的最佳算法。
- **可扩展且极具表达力的发现框架：**该方法的成功关键在于两个方面：一个极具表达力的搜索空间和一个前所未有的规模。其“元网络”通过学习更新目标来定义 RL 规则，这比以往仅调整超参数或单个损失项的方法更为通用。同时，该方法在大规模、多样且复杂的环境（如 Atari, ProcGen, DMLab）中进行元学习，这对于发现一个强大且通用的规则至关重要。
- **涌现出新颖的预测语义和强大的泛化能力：**对 DiscoRL 的分析表明，它学会了进行与传统价值函数不同的新颖预测。这些被发现的预测对于未来的关键事件（如获得高额奖励或策略熵的变化）具有很强的信息量。至关重要的是，这个被发现的规则表现出卓越的泛化能力，在它发现过程中从未见过的挑战性基准（如 ProcGen）上也取得了顶尖性能，证明了其鲁棒性和广泛的适用性。

相关链接

Paper: https://www.nature.com/articles/s41586-025-09761-x

Github： https://github.com/google-deepmind/disco_rl

● 论文二十二

Tongyi DeepResearch Technical Report

日期

2025-10-28

Tongyi DeepResearch Technical Report

Tongyi DeepResearch Team*

Tongyi Lab , Alibaba Group

内容提要

本文介绍了 Tongyi DeepResearch，一个专为长程、深度信息检索研究任务设计的智能体大语言模型。为了激发自主的深度研究能力，Tongyi DeepResearch 通过一个结合了“智能体中训”（agentic mid-training）和“智能体后训”（agentic post-training）的端到端训练框架进行开发，实现了在复杂任务中可扩展的推理和信息检索。论文设计了一个完全自动化、高度可扩展的数据合成流水线，无需昂贵的人工标注即可为所有训练阶段提供支持。Tongyi DeepResearch 模型总参数量为 305 亿，但每个 token 仅激活 33 亿参数，在多个深度研究基准测试（如 Humanity’ s Last Exam, BrowseComp 等）上均取得了 SOTA 性能。

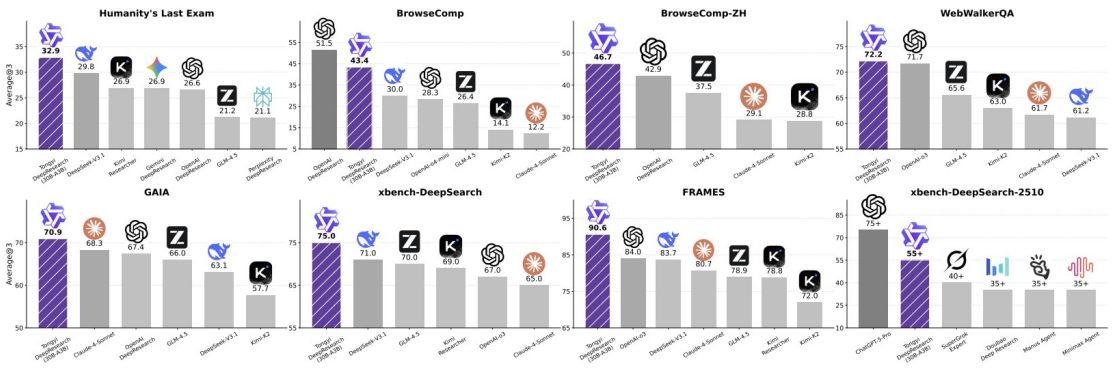


Figure 1: Benchmark performance of Tongyi DeepResearch.

主要亮点

- **创新的端到端智能体训练范式：**Tongyi DeepResearch 提出了一种全新的训练框架，将“智能体中训”和“智能体后训”相结合。中训阶段通过大规模高质量的智能体行为数据，为模型注入强大的智能体归纳偏置；后训阶段则通过强化学习进一步释放其深度研究潜力。这一范式使得模型能够从基础的交互技能逐步进化到高级的自主研究行为，系统性地解决了通用大模型缺乏智能体能力的问题。
- **全自动、可扩展的合成数据引擎：**面对智能体训练数据稀缺且难以标注的挑战，Tongyi DeepResearch 设计了一套完全自动化的数据合成流水线。该流水线能够生成大规模、多样化且质量超越人类水平的智能体轨迹数据，为中训和后训的各个阶段提供定制化、结构化的数据支持，成为驱动智能体能力扩展的核心引擎。
- **卓越的性能与极致的效率：**尽管模型参数规模相对较小（30B 总量，3.3B 激活），Tongyi DeepResearch 在包括 Humanity’s Last Exam、BrowseComp、GAIA、FRAMES 等在内的多个权威深度研究基准测试中均取得了 SOTA（业界最佳）成绩，其性能超越了如 OpenAI-o3、DeepSeek-V3.1 等强大的基线模型。这充分证明了其训练框架和数据策略的有效性，实现了性能与效率的完美平衡。

相关链接

Paper: https://arxiv.org/abs/2510.24701

Github: https://github.com/Alibaba-NLP/DeepResearch

Model: https://huggingface.co/Alibaba-NLP/Tongyi-DeepResearch-30B-A3B

Blog： https://tongyi-agent.github.io/blog/

TECHNICAL UPDATES
技术动态

技术动态一

阿里巴巴发布并开源 Qwen3-VL-30B-A3B-Instruct 与 Thinking

日期

2025-10-04

内容提要

阿里通义千问 Qwen 大模型团队开源 Qwen3-VL-30B-A3B，包含 Instruct 和 Thinking 两个版本。MoE 架构，总参数 30B，激活参数 3B。体积更小，性能依然强劲，仅需激活 30 亿参数，性能媲美 GPT-5-Mini 与 Claude4-Sonnet，在 STEM、视觉问答（VQA）、OCR、视频理解、智能体（Agent）任务等多个领域表现出色。

相关链接

Qwen Chat： https://chat.qwen.ai/?models=qwen3-vl-30b-a3b

Github&Cookbooks： https://github.com/QwenLM/Qwen3-VL/tree/main/cookbooks

API： https://www.alibabacloud.com/help/en/model-studio/models#5540e6e52e1xx

Blog： https://qwen.ai/blog?id=99f0335c4ad9ff6153e517418d48535ab6d8afef&from=research.latest-advancements-list

ModelScope： https://modelscope.cn/collections/Qwen3-VL-5c7a94c8cb144b

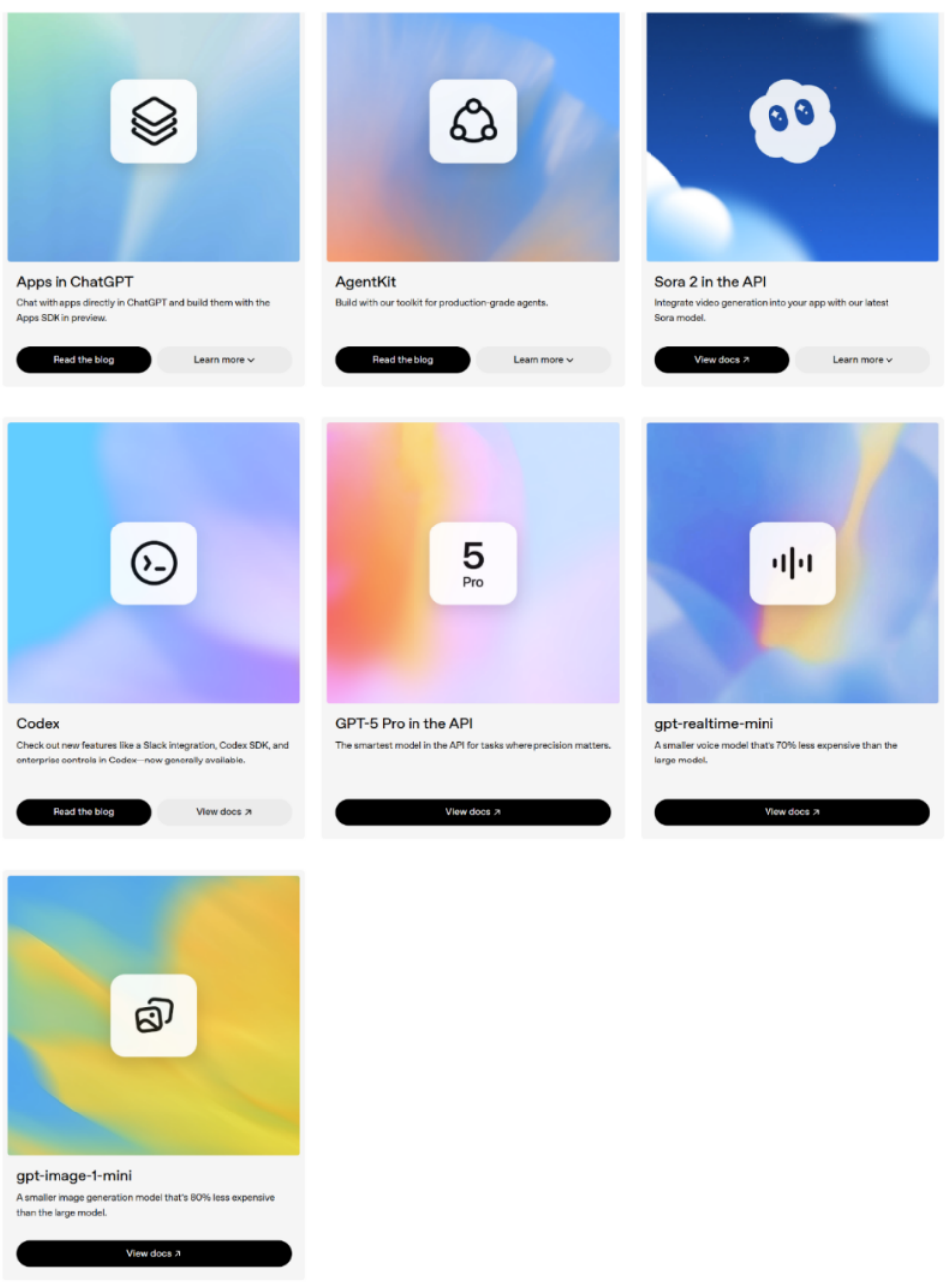
HuggingFace： https://huggingface.co/collections/Qwen/qwen3-vl

• 技术动态二

OpenAI 开发者大会

日期

2025-10-07



内容提要

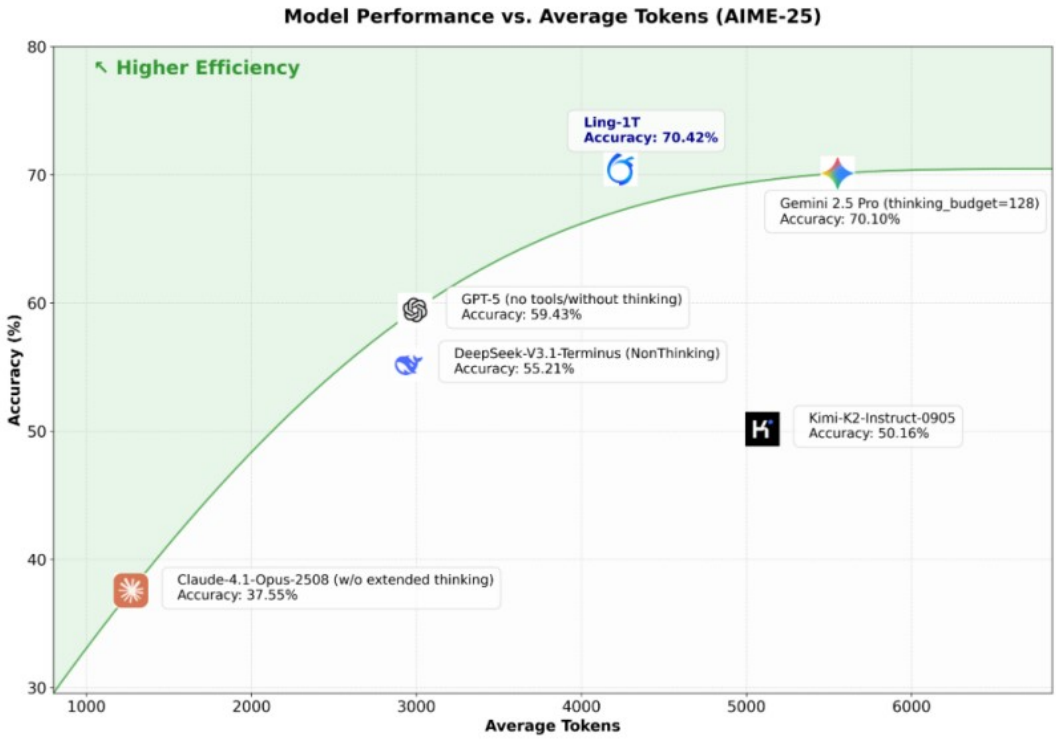
在 OpenAI DevDay 2025 开发者大会上，OpenAI 发布了其首个完整的智能体（Agent）开发平台 AgentKit，旨在大幅简化和加速智能体的构建、部署与优化。该平台通过一套全新的模块化组件实现其核心功能：可视化的 Agent Builder，允许开发者通过拖拽方式设计多智能体工作流并配置开源的 Guardrails 安全护栏；集中的 Connector Registry，用于统一管理数据源与工具连接；以及 ChatKit 工具套件，可将定制化的聊天智能体直接嵌入应用或网站。为进一步强化智能体能力，OpenAI 还推出了强化微调（RFT）功能，允许开发者通过自定义工具调用和评估标准来定制推理模型，该功能已在 o4-mini 上全面开放，并在 GPT-5 上进入私测。与此同时，多个重要产品与模型也迎来更新：Codex 正式版上线，带来了全新的 Slack 集成和 Codex SDK；ChatGPT 通过全新的 Apps SDK 支持直接在聊天界面中运行的可对话应用；并发布了多个新模型 API，包括用于实时交互的 gpt-realtime-mini、原生多模态的 gpt-image-1-mini、备受瞩目的 GPT-5 pro，以及首次通过 API 开放的视频生成模型 Sora 2（提供速度优先的 Sora 2 和画质优先的 Sora 2 Pro 两个版本）。此次系列发布标志着 OpenAI 的战略重心从提供基础模型转向构建一个全面的、端到端的 AI 开发平台，通过提供从可视化编排、数据连接、安全护栏到应用内嵌的全套工具，极大地降低了复杂 AI 应用和智能体的开发门槛，旨在将 AI 能力更深度地集成到各类业务流程与产品体验中。

• 技术动态三

蚂蚁开源万亿参数语言模型 Ling-1T

日期

2025-10-09



内容提要

蚂蚁集团正式发布并开源了其万亿参数级通用语言大模型 Ling-1T，该模型基于其自研的高效 MoE（Mixture of Experts）架构，旨在通过“帕累托改进”实现强推理能力与高计算效率的平衡，突破了超大模型普遍存在的“精确”与“效率”此消彼长的困境。Ling-1T 在编程、数学等高推理密度任务中表现突出，同时具备强大的知识理解、Agent 推理、多轮对话以及工具调用能力，使其不仅能回答问题，还能通过调用外部资源（如联网搜索、数据库查询）来执行复杂的现实世界任务。在多项基准测试中，Ling-1T 刷新了记录：在编程推理 LiveCodeBench 上得分最高，在综合数学测试 Omni-Math 与 UGMathBench 上均突破 74 分；在知识理解维度，其 C-Eval 得分达 92.19，MMLU-Pro 得分 82.04，多项指标领先于 DeepSeek、Kimi 等模型。在 AIME-25 数学竞赛推理测试中，Ling-1T 以 70.42%的准确率与 Gemini-2.5-Pro 并列最高，但其思考路径（token 消耗）更短，展现了更高的推理效率。尽管 Ling-1T 在多项任务中表现优异，但其核心优势在于“大参数储备+小参数激活”的效率范式，即拥有万亿级总参数，但在单次推理中仅激活约 50B 参数，从而在保证能力的同时控制了计算开销。不过，在实际生成任务中（如代码生成），仍可能出现功能性 bug，需要通过迭代优化来提升稳定性。这一成果得益于蚂蚁在数据、架构和训练范式上的多项创新：首先，通过自建 AI Data System 从 40T+语料中提炼出 20T+高推理密度数据；其次，采用三阶段精细化训练路径，并结合自研的 WSM 调度器与 LPO（语言单元策略优化）强化学习方法，使模型在学习逻辑而非仅记忆词元；最后，其原生 FP8 混合精度训练平台为万亿参数的高效训练提供了算力保障。Ling-1T 现已在 Hugging Face 上开源，并提供在线体验。蚂蚁集团同时开放了 ATorch 框架和强化学习工具链，旨在降低开发者和中小企业的参与门槛，共建生态。此举被视为蚂蚁践行“AI 普惠”理念的关键一步，尤其对于金融、医疗等高合规行业，开源的透明性使其能够被审计和深度定制，加速了 AI 从实验室走向产业化应用，最终目标是让 AI 像电力和支付一样无处不在。

相关链接

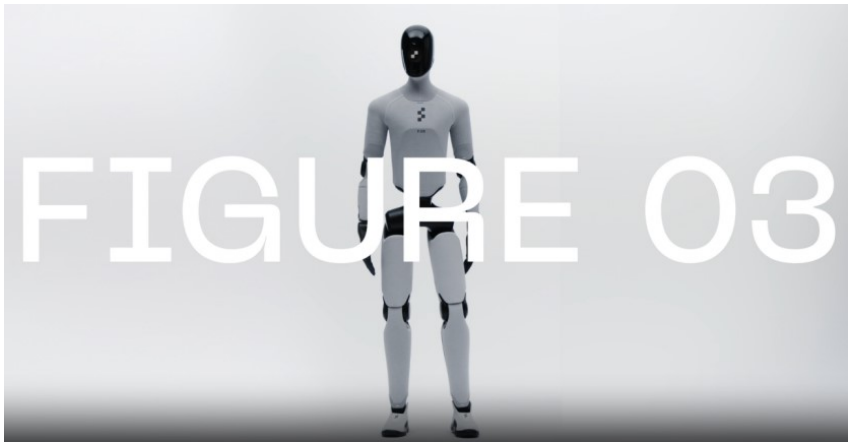
ModelScope：
https://modelscope.cn/models/inclusionAI/Ling-1Thttps://modelscope.cn/models/inclusionAI/Ling-1T
GitHub： https://github.com/inclusionAI/Ling-V2
HuggingFace： https://huggingface.co/inclusionAI/Ling-1T

● 技术动态四

Figure 03 惊天登场

日期

2025-10-09



内容提要

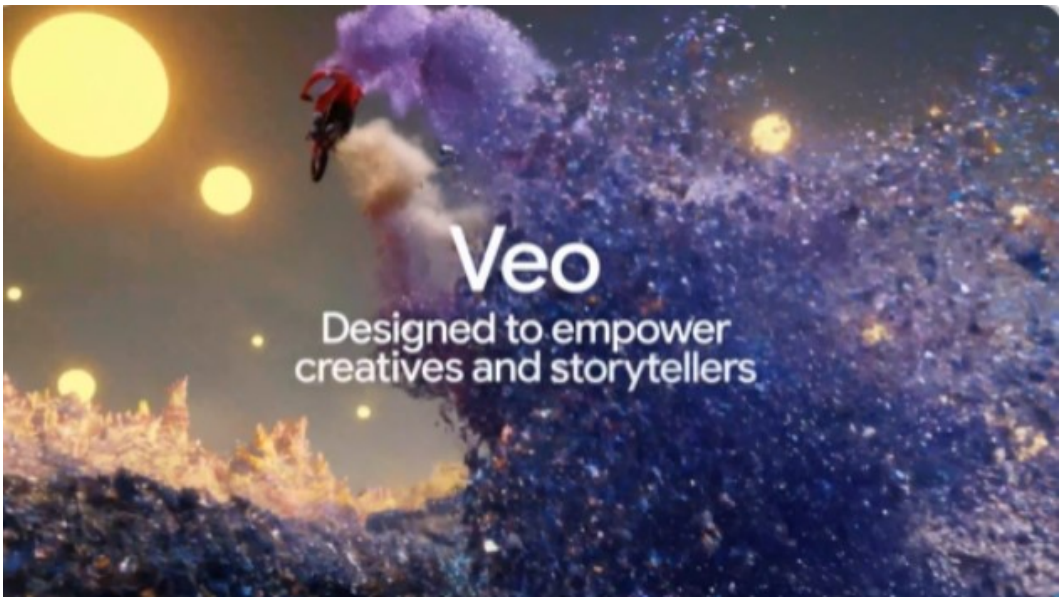
Figure 公司正式发布了其下一代人形机器人 Figure 03，旨在开启通用机器人的规模化时代。该机器人由其自研的视觉-语言-动作模型（VLA）Helix 驱动，使其能够在复杂多变的环境中通过与人类互动来自主学习和执行任务。Figure 03 在硬件上实现了重大突破：其革命性的视觉系统将帧率提升一倍，延迟降低至四分之一，并扩大了 60%的视野，同时在每只手掌心集成了广角摄像头，为抓取任务提供冗余的近距视觉反馈；手部系统实现了质的飞跃，其自主研发的触觉传感器能感知低至 3 克的压力，足以察觉指尖上一枚回形针的重量，从而实现拾取鸡蛋而不捏碎等精细操作。为适应家庭和商业环境，Figure 03 采用了全新的柔性织物外层和多密度泡沫以提升安全性，外层可拆卸水洗，并支持 2kW 的无线感应充电。更重要的是，Figure 03 是首款为大规模制造而设计的机器人，Figure 为此创立了 BotQ 工厂，计划在四年内累计生产十万台。通过 10 Gbps 毫米波数据卸载能力，整个机器人舰队能够上传 TB 级数据进行持续的集体学习，其通用设计使其能无缝应用于家庭（如家务、陪伴）和商业场景（如快递配送、酒店前台、工厂物流），标志着人形机器人在硬件集成、AI 能力和量产规划上迈出了关键一步。

• 技术动态五

谷歌 Veo 3.1 迎来重大更新

日期

2025-10-15



内容提要

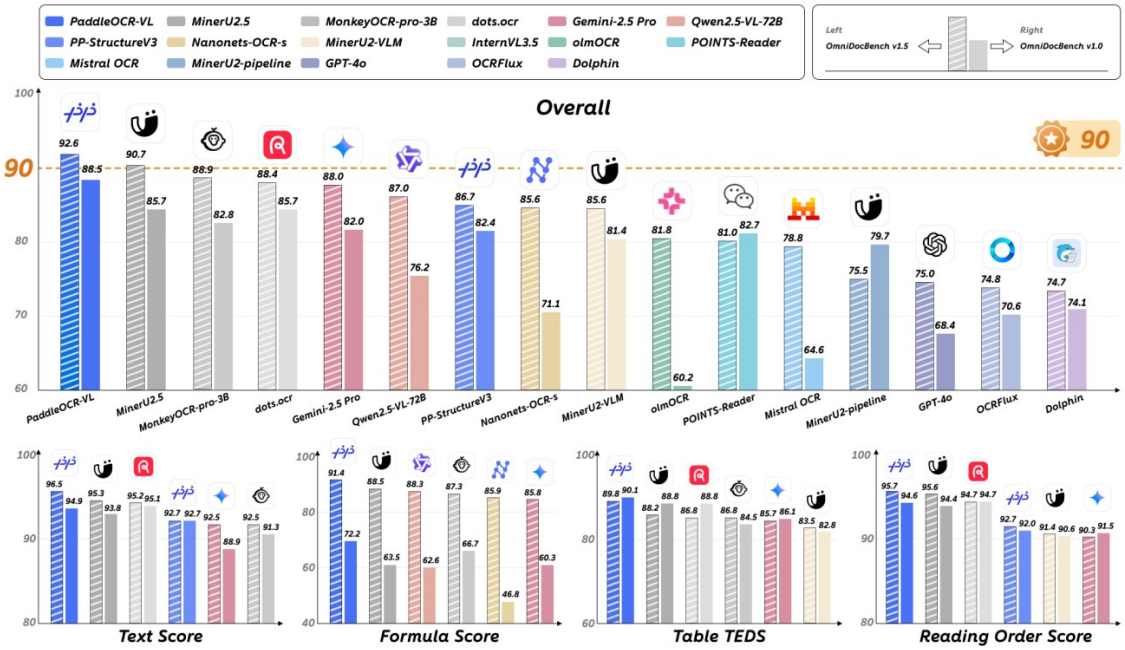
谷歌深夜发布了其最新的 AI 视频生成模型 Veo 3.1，该模型在 Veo 3 的基础上，核心突破在于原生集成了音频生成能力，实现了对对话、环境音效的同步控制，并显著提升了提示词遵循度和质感还原。Veo 3.1 将音频生成功能扩展至“素材生成视频”、“连帧成片”及“延展”等核心应用，允许用户通过文本、单张或多张参考图像（最多三张）生成带有音画同步的视频。该模型支持生成 720p 或 1080p 分辨率、24fps 的视频，基础时长 4 至 8 秒，通过“延展”功能可延长至 148 秒，并能确保延展片段与前一帧的无缝衔接与一致性。同时，它引入了首末帧插值、场景延展以及“插入”和“移除”等更精细的编辑控制能力，增强了叙事和场景的一致性。Veo 3.1 目前通过 AI 电影创作工具 Flow、Gemini API 和 Vertex AI 平台提供，但仍处于预览阶段，且部分高级编辑功能尚未在 API 中完全开放。其定价模式与前代一致，在 Gemini API 的付费层级中提供，标准模型每秒 0.40 美元，快速模型每秒 0.15 美元，且无免费层级。此次更新被视为谷歌在 AI 视频领域对标 OpenAI Sora 的重要举措，通过集成音视频一体化生成，旨在简化企业级内容（如营销、培训视频）的创作流程，减少后期制作需求，标志着 AI 视频生成向更实用、更具控制力的方向迈进。

• 技术动态六

百度开源 PaddleOCR-VL

日期

2025-10-16



内容提要

百度开源 0.9B 参数的模型 PaddleOCR-VL，刷新 OmniBenchDoc 纪录，性能指标目前排名第一，全面超越 GPT-4o、Gemini-2.5 Pro 以及 MinerU2.5、dots.ocr 等前沿模型。推理速度较 MinerU2.5 提升 14.2%，较 dots.ocr 提升 253.01%。支持 109 种语言识别，覆盖中文、英语、法语、日语、俄语、阿拉伯语、西班牙语等多语种场景，大幅提升了模型在多语言文档中的识别处理能力。

相关链接

技术报告：https://arxiv.org/abs/2510.14528

GitHub： https://github.com/PaddlePaddle/PaddleOCR

Models & Demo： https://huggingface.co/PaddlePaddle

• 技术动态七

IROS 圆桌：模型驱动与数据驱动的探讨

日期

2025-10-20

内容提要

在近期杭州举办的 IROS 顶级学术会议中，美团机器人研究院举办了一场圆桌讨论，汇聚了宇树科技王兴兴、浙江大学许超、清华大学许华哲和赵明国等专家。讨论聚焦于机器人智能的核心问题：模型驱动与数据驱动的优劣与融合。专家们从第一性原理、软硬件协同、未来愿景等多角度展开辩论，旨在探索机器人发展的最优路径。

王兴兴指出，当前智能的“第一性原理”仍模糊，类似数据的压缩并非终极答案；未来需像牛顿力学那样总结基本规则。他强调，软件能力越强，硬件要求越低，但现阶段 AI 不足导致硬件一致性需求高。许超提出“牛顿+辛顿”的协同模式，结合物理原理与数据驱动，并设想“云脑”作为付费升级系统。他认为，控制学需从“小脑”转向“大脑”，实现躯体和智能合一。许华哲将智能归结为欲望、先验和经验三要素，主张赋予机器人“学习欲望”，并借鉴生物先验。他以维修工为例，说明经验数据对小众任务的关键性。在模型与数据驱动辩论中，赵明国偏好数据科学，因理论覆盖有限；许华哲以“用脚投票”和“Bitter lesson”强调数据重要性，但最终需理论与数据结合。许超比喻 AI 为“孩子”，自动化为“父”、计算机为“母”、数据为“叔”，呼吁软硬一体。

未来愿景方面，赵明国聚焦机器人足球赛，许华哲设想具好奇心的宇宙探索机器人，王兴兴展望 AGI 作为终极发明。最后，专家鼓励年轻人追寻梦想，投身机器人时代。

相关链接

Blog： https://mp.weixin.qq.com/s/bT7AX63graefS_LUT1rqJw

RESOURCE SHARING

资源分享

资源分享一

《2025 年人工智能现状报告》

类型：讲座报告

内容提要

推理模型成为焦点：OpenAI、Google、Anthropic 和 DeepSeek 在“先思考后回答”推理技术上展开激烈竞争，中国的 R1 系列在部分基准上已超越美方模型，但推理能力仍受“模板匹配”限制。

交互式世界模型兴起：AI 视频从生成短片转向实时交互环境，Genie 3 和 Dreamer 4 等系统为具身智能体训练提供了重要基础。

开源生态崛起但闭源领先：中国 Qwen、Kimi 等开源模型快速进步并赢得开发者青睐，OpenAI 也在 2025 年首次发布开源模型；但顶级性能仍由闭源模型主导。

AI 成为科研伙伴：DeepMind“共同科学家”系统已在药物研发中取得突破，AlphaEvolve 发现了矩阵乘法新算法，AI 正在推动材料科学与生物学的基础创新。

产业进入万亿美元时代：构建前沿智能的成本空前，Stargate 项目投入高达 5000 亿美元；AI 公司收入爆炸式增长，远超传统 SaaS。

英伟达主导算力循环：其市值突破 4 万亿美元，通过投资—采购—再投资形成 GPU 循环生态，巩固芯片霸权。

能源成为新瓶颈：美国数据中心电力缺口预计达 68GW，大量项目因居民反对被搁置，能源供应取代芯片成为 AI 扩张关键限制。

地缘政治加剧 AI 竞赛：美国推行“美国 AI 行动计划”，中国加速自立并拓展“全球南方”影响力，欧洲监管领先但创新滞后，海湾国家以能源打造“主权 AI”中心。

安全形势严峻：AI 网络攻击能力每 5 个月翻倍，实验室强化生物与网络安全防护，但外部安全组织资金不足；模型“伪造对齐”问题凸显，可解释性技术初现突破。

未来趋势值得关注：预测包括中国实验室在主流榜单超越美国、AI 驱动的网络攻击引发国际安全辩论、开源回潮、数据中心邻避主义影响选举，以及 AI 在零售、影视与科研中的颠覆性应用。

资源链接：https://www.stateof.ai/

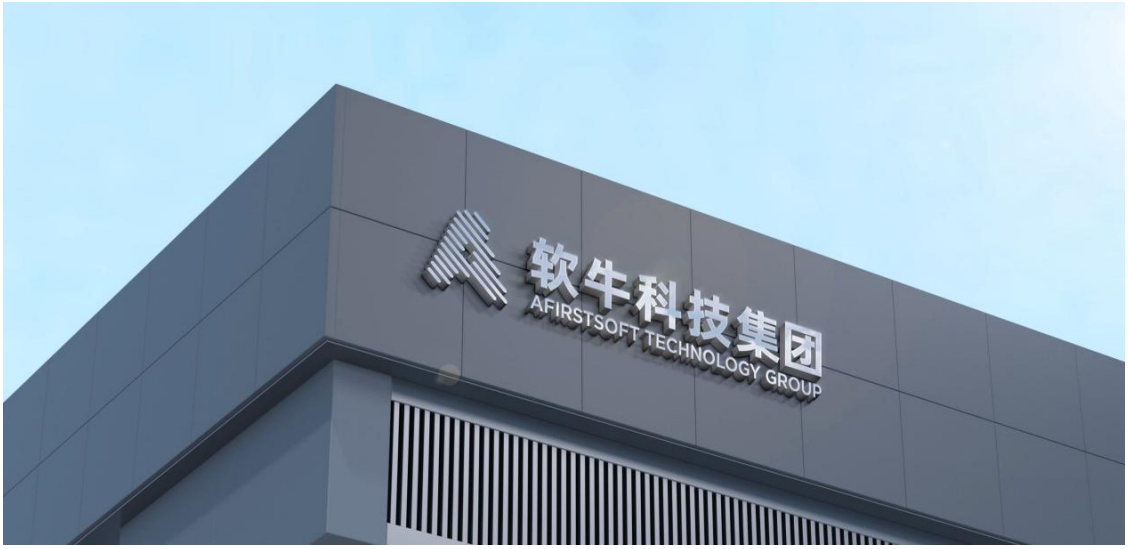
PPT:
https://docs.google.com/presentation/d/1xiLl0VdrlNMAei8pmaX4ojIOfej6lhvZbOIK7Z6C-Go/edit?slide=id.g376435918bd_0_54#slide=id.g376435918bd_0_54

深圳软牛科技集团股份有限公司成立于 2014 年 6 月，作为国家级专精特新“小巨人”企业，始终以“让创意成为现实”为使命，专注于通过人工智能视觉大模型技术驱动全球数字创意生态。

公司深度聚焦图像修复、视频增强、多模态内容生成等视觉大模型核心技术，联合西安电子科技大学、香港中文大学（深圳）等顶尖高校，攻克了视频抠像、画质修复等难题，并与深圳大学共建“人工智能技术联合创新中心”，推动 AI 赋能的图像增强技术规模化落地。

目前，软牛科技能够为广电机构、影视制作、工业检测、农业测报等多个行业提供专业解决方案。从微米级工业精密测量，到食品异物智能检测；从农业害虫 AI 识别防治系统，到 AR 设备巡检实时画面 AI 增强与场景识别；从 VR 全息空间建模助力文博展陈数字化升级，到跨终端智能影像协作软件的全链路画质优化引擎，软牛科技正以技术革新重新定义新质生产力。

在消费级市场，旗下牛小影（https://www.niuxuezhang.cn/video-enhancer.html）、HitPaw 等海内外知名品牌，以“一键式 AI 视觉创意”为核心，为全球超 1 亿用户提供视频修复、老电影 AI 上色等场景化服务。例如，牛小影支持 4K 或 8K 超高清修复技术，帮助家庭用户重现祖辈黑白影像的生动细节；HitPaw Edimakor 通过 AI 语音合成、多语言实时翻译等功能，让自媒体创作者轻松实现跨语言内容生产。



这些创新成果的背后，是公司每年近亿元的研发投入和占比超 50%的顶尖技术团队，以及 CMMI、ISO56005 等国际权威认证体系支撑的硬核实力。立足深圳总部，我们通过北京、新加坡、香港等战略支点构建起辐射全球 200 多个国家和地区的智慧服务网络。作为高新技术企业、深圳市潜在科技独角兽，软牛科技始终以“技术无界”为理念，持续拓展 AIGC+产业应用的边界。

未来，软牛科技将加速推进物理世界与 AI 视觉的无缝衔接，让每个人都能通过指尖的艺术级创作，开启属于自己的数字时代。

人工智能前沿快报

Artificial Intelligence Newslettler

2025 年第 10 期

总第 24 期

广东省图象图形学会

主办 **广东省图象图形学会**

人工智能
前沿快报

2025 年第 10 期
总第 24 期
2025.11.03

声明：

（1）本快报内容提要、文章亮点等部分内容由 AI 生成，不一定准确及全面，文章完整内容及思想应以原文为准。

（2）本文大部分插图来源于相关论文或文章，版权归原作者或原出版社所有。

（3）本快报摘录文章观点不代表编委会及主办方立场。

学会官方公众号

广东省图象图形学会
GDSIG
学术交流、科普宣传、科技成果鉴定