

最近読んだ面白い論文 (AI for Science も含む)

2026/02/19 @slack

tax_free

2026/02/19

会の目的

目的

- ・ AI for Science に使えそうな研究を catch up する
- ・ Slack に論文を貼るだけでなく、集まってチームで議論したい
- ・ 朝にやることで生活リズムを整える (朝活)

1 行ルール

- ・ 論文の網羅紹介はたぶんしません。「どうすれば AI for Science に応用できそうか」「どこがおもしろいのか」など個人の視点を中心に話します。

セクション 1: 最新モデル動向 (Qwen 3.5 & GLM-5)

A: Contribution

- ・ **この紹介の貢献**: 2月に立て続けにリリースされた中国発フロンティア級オープンモデル 2つの技術的な差分と設計思想を比較する
- ・ **組織**: Qwen 3.5 → Alibaba / GLM-5 → Zhipu AI (Z.ai)
- ・ **公開日**: Qwen 3.5: 2026/02/16, GLM-5: 2026/02/11
- ・ **なぜ今?**: Twitter で盛り上がっていた。ベンチマーク性能が proprietary に並ぶと話題。中国勢はアーキテクチャを工夫しているのでテクニカルにおもしろい
- ・ **使えるとしたら?**: 自分たちでホストして LoRA 等でラボ用にカスタム (データ安全性的に自前ホストが必要)。トークン単価が安い (Opus 比 10x 以上) ので雑にたくさん投げる系に向く
- ・ **リンク**: Qwen 3.5: github.com/QwenLM/Qwen3.5 / GLM-5: huggingface.co/zai-org/GLM-5

B: Method

1. **対象 (入力→出力)**: text (+ image/video for Qwen 3.5) → text
2. **キモ (新規点)**:
 - Qwen 3.5: 397B MoE (17B active)。75% の層で attention を **Gated DeltaNet** (線形 attention) に置換 → 準線形スケーリング。Native multimodal
 - GLM-5: 744B MoE (40B active, 256 experts top-8)。MLA + DSA。**Muon Split**: head ごとに独立した行列直交化。全て Huawei Ascend 910B で学習
3. **流れ**:
 - Qwen 3.5: multimodal early fusion pre-training → massive-scale RL
 - GLM-5: 3 段階 pre-training (27T tokens → context 拡張 32K→200K → DSA 適応) → 4 段階 RL (Reasoning → Agentic → General → Distillation)。slime (非同期 RL) + APRIL

4. **評価のしかた**: AIME 2026, SWE-bench Verified, BrowseComp, HLE, MMLU, GPQA Diamond 等

C-1: Results (ベンチマーク比較)

Evidence 1

- 条件: フロンティアモデルとの主要ベンチマーク比較

Benchmark	Qwen 3.5	GLM-5	GPT-5.2	Opus 4.5
AIME 2026	91.3	92.7	96.7	93.3
SWE-bench Verified	76.4	77.8	76.2	80.9
BrowseComp	78.6	75.9	59.2	67.8
Humanity's Last Exam	–	50.4	45.8	43.4
Terminal-Bench 2	52.5	56.2	54.2	59.3

- 意味: 中国発オープンモデルが frontier 級に到達。Qwen 3.5 は agentic search、GLM-5 は knowledge reliability に強い。ただし coding では Opus 4.5 がリード

C-2: Results (設計思想の違い)

Evidence 2

- ・ 条件: アーキテクチャと学習戦略の比較
- ・ **Qwen 3.5**: Gated DeltaNet で **attention そのものを置換** → 長文推論速度に全振り (8.6x faster @32K, 19x @256K)。Multimodal native
- ・ **GLM-5**: MLA + DSA で **attention を効率化** + Muon Split で **optimizer を改良**。slime で非同期 RL。NVIDIA 非依存を達成
- ・ 意味: “attention をどう扱うか”で設計思想が分岐。Qwen=置換路線、GLM=効率化路線。どちらも frontier 級に到達 → attention 設計の正解は一つではない

D-1: Gaps (ベンチマーク ≠ 実用の乖離)

Gap 1

- **Applicability Gap:** ベンチ上では proprietary に並ぶが、GLM-5 の verbosity (他モデル中央値 15M に対し 110M tokens 生成) のように実用では大きな挙動差がありうる
- **Evidence status:** ベンチマークは標準タスク (coding, math, search) が中心。AI4S のような専門的・multi-step なタスクでの挙動は未評価
- **Minimal test:** AI4S 系タスク (例: 論文要約→仮説生成→実験計画) を 2-3 個用意し、open vs proprietary で出力を比較。verbosity, hallucination 率, 指示追従性を定性評価

D-2: Gaps (AI4S ハイエンドタスクの評価不在) & Next Step

Gap 2

- **Applicability Gap:** 既存ベンチ (AIME, SWE-bench 等) は汎用タスク。AI4S で求められる「ドメイン知識の正確さ」「不確実性の表現」などはカバーされていない
- **Evidence status:** FrontierScience (前回紹介) のような science-specific ベンチは存在するが、これらのモデルでの評価結果は未公開
- **Minimal test:** FrontierScience の公開問題や domain-specific な質問セットで open vs proprietary を走らせ、専門家評価で差を定量化

Next step: (i) AI4S 系タスクでの簡易比較実験を設計 (ii) verbosity/cost 比を含む実用指標の評価フレームワーク検討 (iii) FrontierScience ベンチで走らせてみる

セクション 2: DeepSeek-OCR 2

A: Contribution

- ・ **この論文の貢献:** CLIP ViT を LLM アーキテクチャ (Qwen2-0.5B) に置換した vision encoder で、文書の意味的な読み順を自律学習する Visual Causal Flow を提案。3B params (500M active) で学術論文 OCR の SoTA
- ・ **著者/組織:** DeepSeek AI (Haoran Wei et al.)
- ・ **公開日:** 2026/01/27
- ・ **なぜ今?:** vision を使った AI4S に興味がある。科学文書は複雑レイアウト (multi-column, 数式, 表) が多く Visual Causal Flow が効くドメイン
- ・ **使えるとしたら?:** 論文の高速な文字起こし→埋め込みパイプライン (16GB GPU 単体で動く)。bidirectional encoder に causal query を追加する設計発想を別ドメインに応用
- ・ **論文リンク:** <https://arxiv.org/abs/2601.20552>

B: Method

1. **対象 (入力→出力)**: 文書画像 (images/PDFs/tables/数式) → Markdown/HTML/テキスト (レイアウト保持)
2. **キモ (新規点)**:
 - DeepEncoder V2: CLIP ViT を Qwen2-0.5B で置換。dual-stream attention (visual tokens=双方向, learnable queries=causal) で読み順を自律学習
 - Visual Causal Flow: raster scan (左上→右下固定) ではなく、query が「次にどこを読むか」を学習。reorder された query tokens のみ decoder に渡す
3. **流れ**: SAM-base tokenizer (16x 圧縮) → DeepEncoder V2 (dual-stream) → query tokens → DeepSeek-3B MoE。3 段階学習: encoder pretraining (100M pairs, 160 A100) → joint training → decoder-only (encoder 凍結, 速度 2 倍)

4. **評価のしかた**: OmniDocBench v1.5 (edit distance, Formula CDM, Table TEDS, reading order)

C-1: Results (OmniDocBench v1.5 SoTA)

Evidence 1

- 条件: OmniDocBench v1.5 で他モデルと比較

Model	Overall	Formula CDM	Table TEDS
DeepSeek-OCR 2	91.09	90.31	87.75
Gemini-3 Pro	88.03	–	–
Qwen3-VL-235B	89.15	–	–
DeepSeek-OCR 1	87.36	84.14	85.25

- 意味: 3B (500M active) という軽量モデルが 235B クラスを上回る。学術論文 edit distance 0.013、研究レポート 0.008 と科学文書が最得意カテゴリ

C-2: Results (Visual Causal Flow の効果)

Evidence 2

- ・ 条件: OCR 1 (raster scan) vs OCR 2 (Visual Causal Flow) を文書タイプ別に比較
- ・ 結論: colorful textbook 0.130→0.053、handwritten 0.145→0.068、PowerPoint 0.052→0.031 と複雑レイアウトで大幅改善。ただし新聞 0.131→0.139 と悪化 (学習データ不足)
- ・ 意味: causal flow が複雑レイアウトの読み順改善に効いている。一方でデータ不足カテゴリには効かない → 学習データ設計が重要

Gap 1

- **Applicability Gap**: OCR (テキスト抽出) は強力だが、図・ダイアグラムの**意味的な理解**は対象外。AI4S では図の解釈や数値の妥当性チェックも必要
- **Evidence status**: OmniDocBench は文字列の正確さが中心。意味的な正しさ (例: 数式の物理的妥当性) は未評価
- **Minimal test**: figure caption と図の対応、表中の数値の単位・桁の整合性チェックを行い、semantic error の割合を測る

D-2: Gaps (single-pass 処理の限界) & Next Step

Gap 2

- **Applicability Gap**: iterative refinement がなく、1 回で読み取れなかった箇所を修正できない。dense page では visual token 上限 (1,120) が制約
- **Evidence status**: 新聞のような高密度レイアウトでは実際に悪化。著者も将来課題として言及
- **Minimal test**: 科学論文の supplementary (dense tables, multi-panel figures) で処理し、エラーパターンを分析

Next step: (i) 手元の論文 PDF で OCR 精度確認 (数式・表) (ii) 16GB GPU でデプロイ試行 (iii) OCR→LLM パイプライン (文字起こし→要約→Q&A) の構築検討

セクション 3: Reinforced Attention Learning (RAL)

A: Contribution

- ・ **この論文の貢献:** マルチモーダル LLM の RL post-training で、出力トークンではなく attention 分布を policy-gradient で直接最適化。GRPO が画像理解を劣化させる問題を解決
- ・ **著者/組織:** UC Davis, Google DeepMind, Princeton (Bangzheng Li, Jianmo Ni et al.)
- ・ **公開日:** 2026/02/04
- ・ **なぜ今?:** AI4S 文脈での VQA に興味がある。DeepSeek-OCR 2 と組み合わせれば multi-panel や複雑な図の読解に使いそう。TSUBAME の H100 で~20,000 円で再現可能
- ・ **使えるとしたら?:** 「どこを見るか」が重要な複雑な VQA 等の multimodal 系タスク。科学画像など情報密度の高い入力で特に有効
- ・ **論文リンク:** <https://arxiv.org/abs/2602.04884>

B: Method

1. **対象 (入力→出力)**: マルチモーダル入力 (image/video + text) → テキスト回答。
内部的に attention 分布を最適化
2. **キモ (新規点)**:
 - Attention 分布を RL policy 化: 最終層の全 head 平均 attention weight を policy とし、正解時→JSD 最小化 (パターン強化)、不正解時→JSD 最大化 (パターン回避)
 - $L_{\text{total}} = L_{\text{RL}} (\text{GRPO}) + \lambda * L_{\text{AttnRL}}$ 。token-level の $O(T)$ に対し attention-level は $O(T^2)$ の勾配 → denser supervision
 - On-Policy Attention Distillation: 32B→7B へ attention 分布も蒸留 (「どこを見るか」も模倣)
3. **流れ**: SFT (165k pairs, 8xH100 ~10h) → RL with RAL (51.2k, G=8 rollouts, 8xH100 ~120h)。visual encoder 凍結、LM backbone のみ更新。
報酬: $0.9 * \text{accuracy} + 0.1 * \text{format}$

4. **評価のしかた**: image 8 ベンチ (V^* , ChartQA, MME 等) + video 7 ベンチ (LongVideoBench, NExTQA 等)

C-1: Results (GRPO の画像劣化を RAL が解決)

Evidence 1

- 条件: Qwen-2.5-VL-7B に GRPO vs RAL を適用し比較

Image Bench	Base	GRPO	RAL
V*	70.7	68.6 ↓	73.3
ChartQA	84.0	81.7 ↓	86.4
MME	2309	2259 ↓	2353
MuirBench	44.9	43.9 ↓	47.4

- 意味: GRPO は画像ベンチ 8 つ中 5 つで base より**悪化**。RAL は 8 つ全てで改善。RL post-training のマルチモーダル適用時に「視覚能力劣化」が起きる根本問題を attention 最適化が解決

C-2: Results (visual token 数が増えるほど優位性拡大)

Evidence 2

- ・ 条件: V*Bench で画像解像度 (visual token 数) を 512→2048 に変化させて RAL vs GRPO を比較
- ・ 結論: 512 tokens で+1.6 の差が、2048 tokens で**+6.3** に拡大。RAL-zero (CoT なし) でも GRPO+CoT に勝利
- ・ 意味: visual 情報が多い=attention の割り当てが重要なタスクほど RAL が効く。科学画像・実験動画など情報密度の高い入力に有望。CoT なしでも効く → 「見る能力」が本質的なボトルネック

Gap 1

- **Applicability Gap:** 実験は Qwen-2.5-VL (7B/32B) のみ。他アーキテクチャでの汎化は未確認。最終層 attention head 平均という設計がアーキテクチャ依存の可能性
- **Evidence status:** 単一モデルファミリーでの一貫した改善は示されている
- **Minimal test:** InternVL や LLaVA-NeXT など別アーキテクチャで再現するか確認

D-2: Gaps (exact-match 報酬の限界) & Next Step

Gap 2

- **Applicability Gap:** 報酬が exact match (1 or 0) のみ。AI4S での科学的推論は open-ended で二値評価できない
- **Evidence status:** 選択式・短答式タスクでの改善のみ。生成型タスクでの効果は未検証
- **Minimal test:** LLM-as-a-judge 等の連続報酬で RAL が機能するか、science QA タスクで試す

Next step: (i) science VQA で GRPO vs RAL の差を確認 (ii) attention 可視化で「何を見ているか」を定性分析 (論文にないので自分でやる価値あり) (iii) 科学画像 (顕微鏡画像、スペクトル等) での効果を検証

Appendix

A: Glossary

- **MoE**: Mixture of Experts の略。全パラメータのうち一部のみを推論時に活性化
する手法 (例: Qwen 3.5 は 397B 中 17B のみ active)
- **SFT**: Supervised Fine-Tuning の略。ラベル付きデータで教師あり学習する
post-training の最初のステップ。人手やモデルが作った (入力, 出力) ペアで学習
する
- **Pre-training / Mid-training / Post-training**: LLM の学習 3 段階。Pre-
training: 大量テキストで次トークン予測 (基礎能力獲得)。Mid-training: context
拡張やドメインデータの追加学習。Post-training: SFT→RL で指示追従や品質を
磨く。GLM-5 は 3 段階 pre-training + 4 段階 post-training という多段構成

Q1 (DeepSeek-OCR 2): 新聞が苦手なのはアーキテクチャの問題？ データ不足？

- ・ A: 主にデータ不足。学習データに新聞は~250K サンプルしかなく、著者も“insufficient training data”を原因として挙げている。ただしアーキテクチャの制約も間接的に効いている: visual token 上限が 1,120 なので情報密度の高い新聞ではトークン不足になりうる。また single-pass 処理のため読み取れなかった箇所を再検査できない。著者も将来課題として“longer causal flow tokens for multiple re-examinations”に言及。データ増で改善するはずだが、token 上限で頭打ちになるかは未検証。