

DeepSeek-V4-Pro-0813：圧倒的なコスト性能を誇るAIエージェントの新基準

技術仕様と圧倒的なエージェント性能



1.6兆パラメータのMoE構造と1Mコンテキスト

490億のアクティブパラメータと、100万トークンの広大な処理空間を両立。



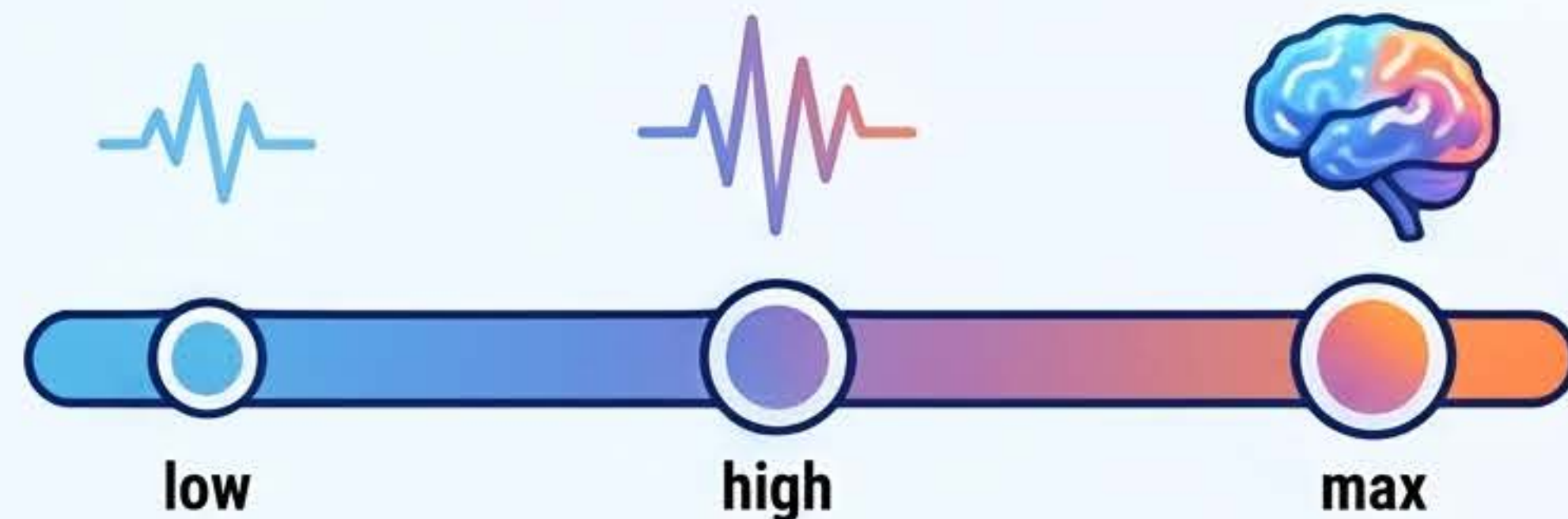
Terminal-Bench 2.1
(V4-Pro 0813)



Claude Code

ターミナル操作精度が約21.9%向上

Terminal-Bench 2.1で87.9点を記録し、Claude Code等の競合を凌駕。



3段階の思考強度制御（Thinking Effort Control）

タスクの複雑さに応じてlow/high/maxの推論リソースを動的に調整可能。



2026年8月13日に正式版（GA）がリリースされたDeepSeek-V4-Pro-0813は、高度な推論制御と圧倒的な低価格を実現し、特に自律型エージェントやコーディングタスクにおいてフロンティア級の性能を発揮します。

主要エージェント・ベンチマーク進化と競合比較

ベンチマーク指標	Preview版 (4月)	V4-Pro 0813 (GA)	比較対象 (Claude Fable 5等)
Terminal-Bench 2.1	72.1	87.9	88.0 (Fable 5) 83.5 (Claude Code)
CyberGym (セキュリティ)	52.7%	83.3%	—
DeepSWE (ソフトウェア開発)	12.8%	62.7%	—

革命的なコスト構造と新料金モデル



競合モデル
(Claude Fable 5)

出力単価は競合モデルの「約57分の1」

Claude Fable Sと比較し、極めて低いトークン単価でフロンティア級性能を提供。



DeepSeek-V4-Pro-0813



コンテキストキャッシュによるコスト最適化

キャッシュヒット時の入力単価は100万トークンあたり\$0.003625と激減。



ピーク オフピーク

業界初、時間帯別料金制（ピーク/オフピーク）

オフピーク時にタスクをスケジューリングすることで、コストをさらに半減可能。