

# MinerU2.5-Pro: A 1.2B Model Beats 200x Larger Rivals by Fixing the Data, Not the Architecture

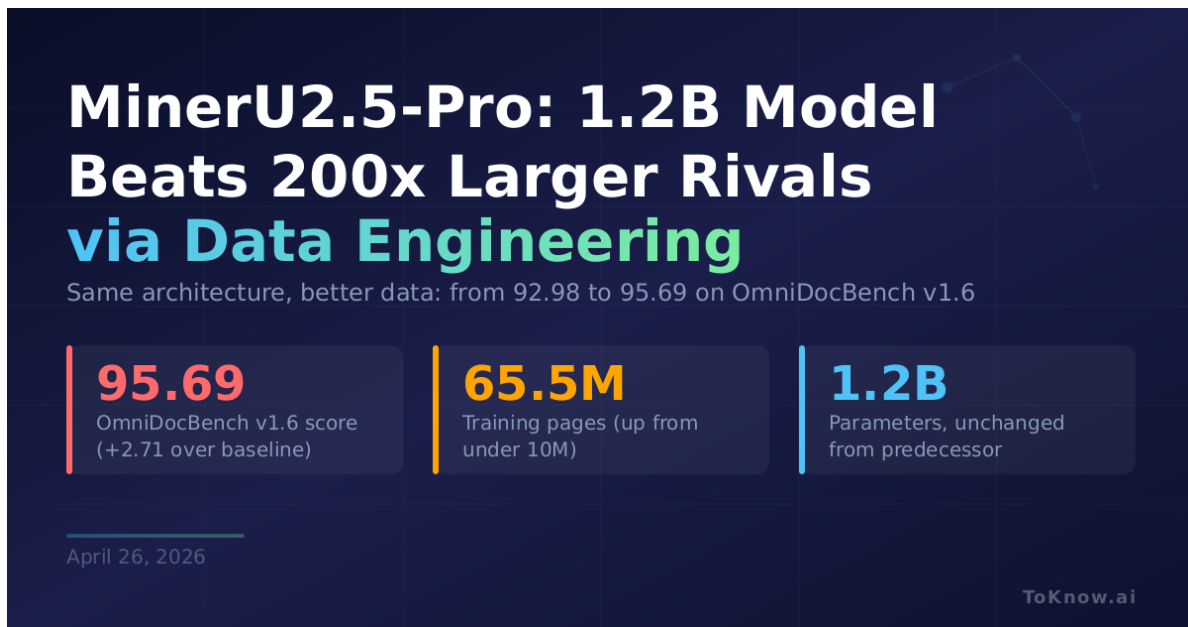
Kabui, Charles

2026-04-26

---

[Read at ToKnow.ai](#)

---



OpenDataLab released [MinerU2.5-Pro](#), a document parsing model that converts PDFs into structured Markdown. It keeps the exact same 1.2B-parameter architecture as its predecessor and improves the overall score on [OmniDocBench v1.6](#) from 92.98 to **95.69**, beating every

other system tested, including Gemini 3 Pro, Qwen3-VL-235B, and specialized OCR models with 200x more parameters. All gains came from a three-component Data Engine. Diversity-and-Difficulty-Aware Sampling expanded training data from under 10M to 65.5M pages while correcting long-tail distribution bias. Cross-Model Consistency Verification ran three different models on each sample and used their agreement to classify difficulty and generate reliable annotations. A Judge-and-Refine pipeline rendered LaTeX and HTML outputs back into images, then visually compared them to the originals to catch structural errors that text-only checks miss. A three-stage training progression (large-scale pretraining, hard-sample fine-tuning, [GRPO](#) alignment) matched data quality tiers to training phases.

This gives you production-grade parsing of tables, formulas, charts, and multi-column layouts from a model that runs on 4GB of VRAM. For legal, financial, and medical pipelines processing millions of pages, that means enterprise-quality extraction without enterprise-scale GPUs. MinerU now supports PDF, DOCX, PPTX, and XLSX inputs and recently moved to an [Apache 2.0-based license](#), removing a major barrier to commercial adoption.

State-of-the-art models across wildly different architectures fail on the same hard samples, which points to training data as the shared bottleneck. Better annotation pipelines, not bigger models, are the lever. [MinerU-Diffusion](#), from the same team, recently attacked decoding speed; MinerU2.5-Pro attacks data quality. Together they suggest document parsing is entering a phase where infrastructure matters more than architecture.

Sources:

- [MinerU2.5-Pro Paper \(arXiv\)](#)
- [MinerU2.5-Pro Model \(HuggingFace\)](#)
- [MinerU GitHub Repository](#)
- [OmniDocBench Evaluation Framework](#)

---

*Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. **Read more:** [/terms-of-service](#)*