

# OvisOCR2: A 0.8B Model Parses Whole Pages in One Pass

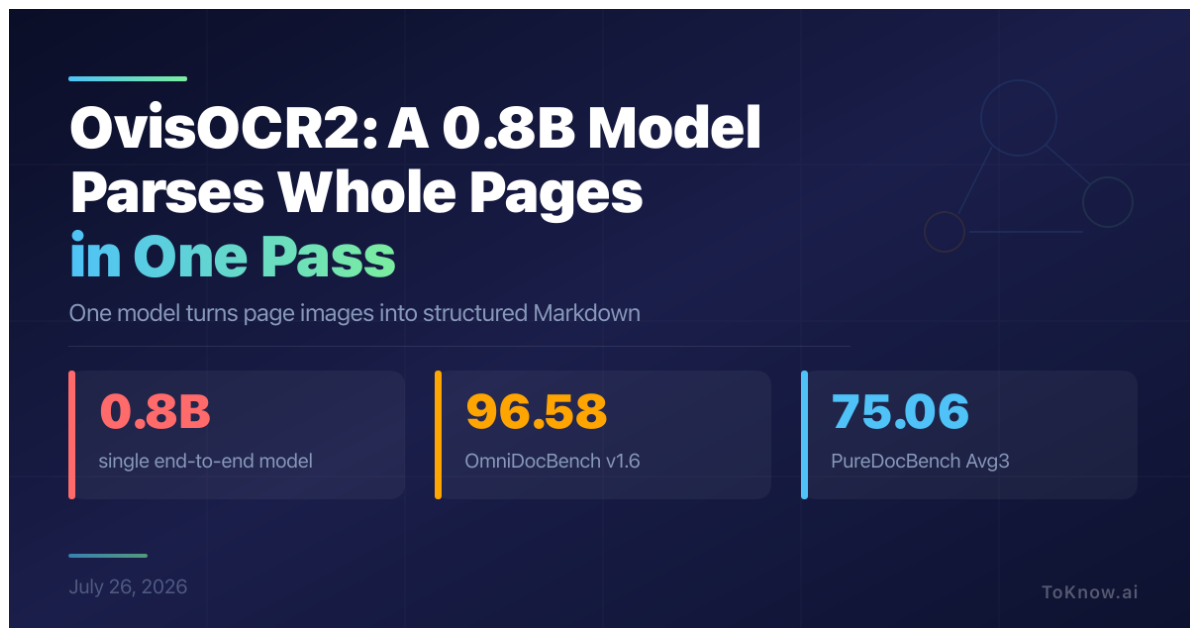
Kabui, Charles

2026-07-26

---

 Read at [ToKnow.ai](https://ToKnow.ai)

---



OvisOCR2 is a 0.8B model that reads a full document-page image and writes its text, formulas, tables, reading order, and visual regions as Markdown in one pass. Most leading parsers first detect page regions, then send each crop to a specialist model. OvisOCR2 skips that chain. Its training mixed filtered annotations from real documents with synthetic pages whose images and correct Markdown came from the same HTML source. The team trained a larger 4B

model with reinforcement learning, then transferred that behavior into the smaller model. In the authors' [OmniDocBench v1.6 results](#), OvisOCR2 scored 96.58, ahead of the pipeline-based PaddleOCR-VL-1.6 at 96.33 and MinerU2.5-Pro at 95.75. It also reached 75.06 across the three [PureDocBench](#) test settings.

One model is easier to deploy than separate layout and recognition stages, and it avoids a common pipeline failure: if the first stage misses a table, later stages cannot recover it. The downloadable [Apache 2.0 model](#) runs through vLLM and supports outputs up to 16,384 tokens, which suits dense pages. The lead is narrow, though, and the paper reports neither latency nor memory use, so it does not prove lower operating cost. Its 66.56 score on PureDocBench's recaptured-image test also trails much larger general models. OvisOCR2 is strongest as a simpler parsing option, not evidence that messy scans are solved.

Read More: [PaddleOCR-VL-1.6: A 0.9B Document Parser That Beats Models Hundreds of Times Its Size](#)

Sources:

- [OvisOCR2 Technical Report](#)
- [OvisOCR2 Model Card and Inference Guide](#)
- [OmniDocBench Evaluation Framework](#)
- [PureDocBench Paper](#)

---

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