

Qwen3.8-27B: A 27B Open Model That Runs Agent Work on One GPU

Kabui, Charles

2026-08-28

[Read at ToKnow.ai](#)



Alibaba released Qwen3.8-27B on August 14 under the permissive Apache 2.0 license, a dense 27-billion-parameter model that natively understands text, images, and video while carrying multi-step agent tasks to completion. It is the compact member of the Qwen3.8 family, built on the Qwen3.5 architecture with hybrid attention, multi-token prediction, and a 262,144-token context window that stretches to 1 million. On Alibaba’s own benchmarks it scores 84.3 on

OSWorld-Verified, where an agent drives a desktop, 61.7 on SWE-bench Pro, and 90.3 on LiveCodeBench v6. Independent testing from [Artificial Analysis](#) puts its Intelligence Index at 52, matching GPT-5.6 Luna at maximum reasoning.

The part developers actually got excited about is the size. A 4-bit build is about 17 GB, so it runs on a single 24 GB GPU such as an RTX 4090, or a Mac with 24 GB of unified memory. Simon Willison tested that build on an M5 Max MacBook Pro and an NVIDIA DGX Spark, using it to write code, interpret images, and drive a coding-agent loop, all offline. It passed 3 million Hugging Face downloads within three days.

There is a real trade-off: thinking mode defaults to extra high, which makes the model slow and wordy, and the headline scores are Alibaba's own until more independent results land. But a model you can run on hardware you own now scores near cloud-only frontier models from a few months ago, the same argument for keeping a [local hedge](#) when a hosted model can vanish.

Read More: the previous generation, [Qwen3.6-27B](#), also runs locally on 18 GB.

Sources:

- [Qwen3.8-27B model card \(Hugging Face\)](#)
- [Qwen3.8-27B runs frontier-class coding agents and reasoning locally \(VentureBeat\)](#)
- [Qwen 3.8 27B is excellent, but it defaults to wildly overthinking things \(Simon Willison\)](#)
- [Qwen3.8 27B intelligence, performance and price analysis \(Artificial Analysis\)](#)
- [Qwen3.8 - How to Run Locally \(Unsloth docs\)](#)

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