

FEVER

TECHNICAL SPECIFICATION

Fourier Encoder for Visual Embedding and Retrieval

Self organizing media files.

For both you and your customers.

FEVER is a bespoke, ultra-efficient model from Lowdown Labs, set into a full multimedia multi-modal retrieval product. It runs inside your VPC, indexes your object storage buckets, and can provide full text -> multimedia search functionality without endless configurations and tuning. Nothing leaves your account, no billing per call. You can use our packaged API in your own application at scale. Customers don't have to organize the file pile, neither do you. Build nothing, turnkey.

AT A GLANCE

Deploy CloudFormation
Peak Ingest . . ~2,000 items/sec
QPS ~5,000 @ 16 core
Compute . . . CPU native, GPU optional
Database . . . Built on PG
Interfaces . . . REST, MCP, console
Egress None required
TCO Est. \$8k USD/mo

THROUGHPUT AND COST, MEASURED

Embed **51** images/sec on one CPU node, **2,666** on one GPU. Query 2.9/sec on CPU and 226/sec on GPU.

(Hardware: c7i.4xlarge int8, g6e.xlarge L40S fp16)

COST TO INDEX A CORPUS, ORDER OF MAGNITUDE. EXTRAPOLATED FROM ONE MEASURED RUN ON 10,000 EXAMPLE IMAGES.

CORPUS	TIME	COST
10 thousand	~2 minutes	~\$0.03
100 thousand	minutes	cents
1 million	hours	dollars
1 billion	days	thousands of dollars

THROUGHPUT BY MEDIA TYPE ON GPU

Images ~2,500 img/sec on GPU
Video ~400 clips/sec
Documents / PDF . . ~1,500 pages/sec
Audio ~50x realtime
Peak ingest ~2,000 items/sec sustained, cross types

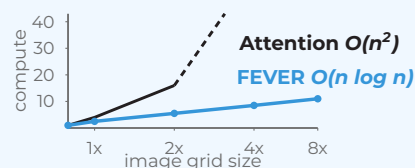
A billion images indexed for a few thousand dollars of compute. **No per-image fee.**

Order of magnitude estimates, extrapolated from 10k image run, not a performance commitment or guarantee.

NOTHING ELSE RUNS ON CHEAP HARDWARE?

- Attention compares every patch to every other patch. $O(n^2)$.
- That is largely why everyone needs GPUs and can't get them.
- FEVER operates in the **Fourier domain** instead. $O(n \log n)$.
- Our CPU numbers can therefore exist.

SCALING



A FULL SYSTEM, NOT JUST A MODEL

Competitors ship 1 piece, DB **or** model. You figure it out from there. **FEVER gives you the whole stack.**

	FEVER	SigLIP	jina-clip-v2
	<i>this appliance</i>	<i>embedding model</i>	<i>model API</i>
Flickr30K Recall@1	81.4	75.6	87.3
Cross resolution recovery	97.8 R@1		
Runs on.	your choice, your VPC	GPU	GPU or metered API

The full pipeline measured above was scored on our patent pending low resolution / high accuracy indexing system, where JINA and SigLIP are slower accepting only **one** resolution, with higher latency and higher storage cost. Neither of the others ships with integrated reranker, a dedup system, or is fully multimodal. **You would otherwise build all of it.**

FEVER numbers measured on the appliance: embed plus the bundled reranker, and the 64 px cross resolution eval.

WHAT FEVER HELPS WITH

- **Easy Training Data Curation.** Train your own AI faster.
- **Natural Language Search.** Ship that feature today.
- **Fight Fraud.** Find dupe content and the uploader.
- **Compliance Checks.** Know what's in your buckets.
- **Recommendations.** Recommend by similarity search.
- **Agentic Search.** Bundled MCP and REST API.

MODEL CARD	COMPONENT	SPECIFICATION
	Image encoder	FEVER, 342M parameters. Proprietary Fourier operator mixing with local attention.
	Text encoder	SigLIP2 so400m text tower (Apache 2.0), aligned to the FEVER image space.
	Reranker	BLIP ITM (BSD 3 Clause), int8 QAT, full 577 token vision sequence.
	Runtime	Custom CPU optimized server. Torch+CUDA bundled for flexible compute setups.

DEPLOYMENT AND SECURITY

- Your VPC, your subnets, your CIDR. Four inputs.
- Easy deploy template, but totally in your control after.
- Dependable scaling and security of Postgres
- Telemetry ready to link with your observability stack.
- No callbacks. No outbound path to reach our models.
- One click deploy takes roughly 30 minutes.

WHAT YOU ARE ACTUALLY BUYING

- No model to choose. No database to configure. No data protection review. Turnkey "sort my enterprise files" system.
- Urgent customer support is included as per listed terms on the procurement platform.
- Your success is our success - The Lowdown Labs Way.

AVAILABILITY

AWS Marketplace.
(GCP Soon)

Container product, annual contract.
Private offers available.

dev@gimmelowdown.com