

github.com/SwayStar123

Researcher at Luma, working on architectures, objectives, scaling laws for generative models.
Previously research at Bria AI, DSL & assembly work at Fuel Labs.

WORK

Luma Labs — Researcher, Pretraining & Scaling

Feb 2026 – Present

- Co-author of *Abra*, a compute-optimal scaling law study for text-to-image diffusion spanning 10^{19} to 10^{22} FLOPs.
- Pretraining and fundamental research on the architectures behind Luma's next generation of generative models.
- Distilling large-scale video and image models - Step distillation and Parameter Distillation

Bria AI — Research Intern

Aug 2025 – Dec 2025

- Trained FIBO-lite single-handedly, the distilled version of the 8B-parameter FIBO.
- Implemented a wide range of research papers for better sampling quality and speed.

Fuel Labs — Software Engineer

Jun 2022 – Apr 2025

- Hand-wrote assembly to optimize standard library functions.
- Worked on the Sway language core and standard library, helping shape how the language is used today; audited ecosystem projects' code.

PAPERS

Abra: Scaling Diffusion Image Training

arxiv.org/abs/2608.17286

2026. Diffusion models scale as predictably as language models, but compute optimality sits near 200 image tokens per parameter, ten times the Chinchilla prescription.

SpeedrunDiT

arxiv.org/abs/2512.12386

Solo author, 2025. Stacks techniques to reach 360× faster training than baseline.

PROJECTS

Reimei

github.com/SwayStar123/reimei

Custom diffusion transformer architecture, folding in numerous papers for efficiency and prompt adherence.

Language Speedrun

languagespeedrun.com

Language-learning site built around SRS and sentence mining; currently using it to learn Chinese.

Paper implementations & replications

github.com/SwayStar123

HDiT · Simple-Relativistic-GAN-Loss · DC_AE · microdiffusion · REPA-Contrastive-Flow-Matching

Samples from two of my own models: *Reimei*, a custom diffusion transformer, and *SpeedrunDiT*, the 360× faster training recipe.



Fig. 1 · Reimei, text-conditioned scene generation.

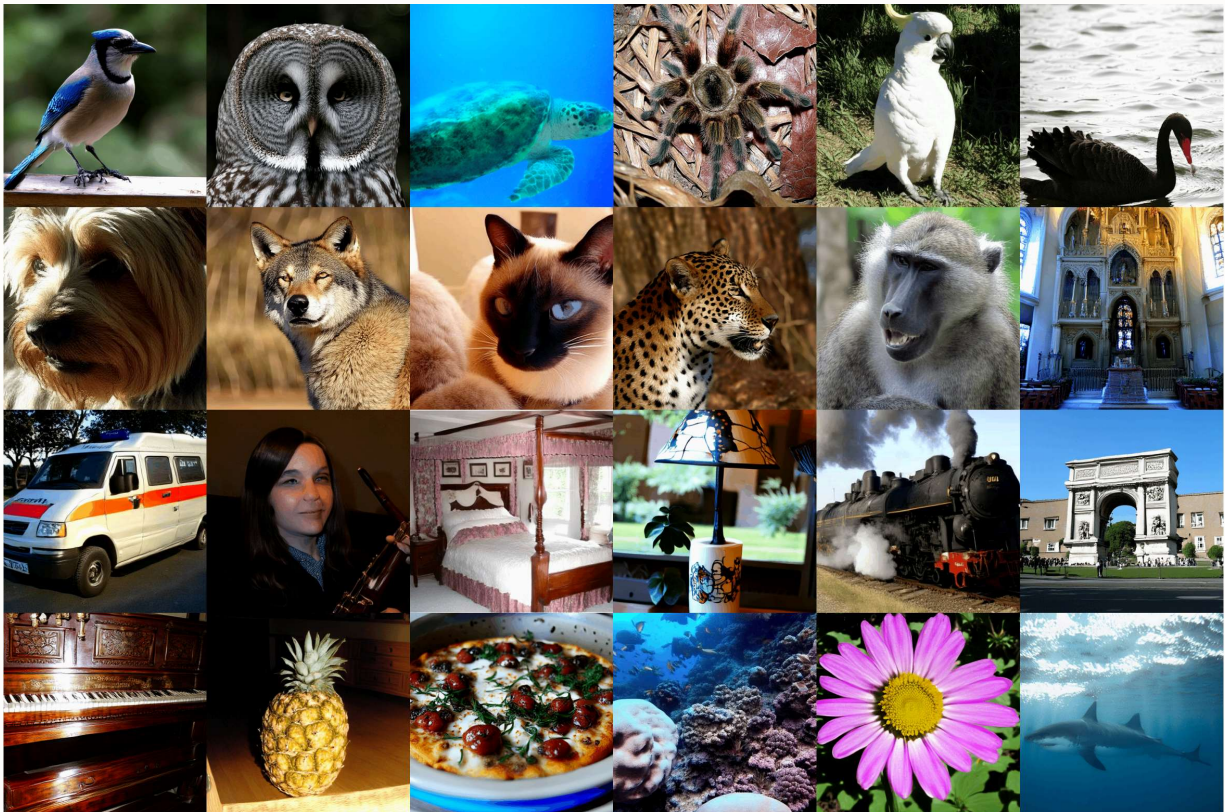


Fig. 2 · SpeedrunDiT, class-conditioned samples.