Lucas Image 2.1 Turbo: three-component inference package

Lucas van de weerd | Lucasgitguy-co | 7 November 2026 | Version 1.0

Repository: Lucasgitguy-co/Lucas-Image-2.1-Turbo-Q2K-SakuraEncoder

Reviewed revision: c8450700eb1d684d92e8c4b02767c11651bbca60

Abstract

This report documents an inference package assembled from existing diffusion, text-encoder and VAE artifacts. The publisher's contribution is packaging, provenance, checksums and usage documentation; the records do not describe new fine-tuning or new quantization by this publisher.

1. Scope and related work

This technical report reconstructs the release from pinned repository artifacts and the publisher's earlier portfolio report. It is an AI-assisted documentation synthesis, not an independent rerun or a peer-reviewed publication. Reported measurements belong to the original release records. [1, 2, 3, 4]

The earlier portfolio report provides the organizing distinction between training, model modification, conversion and packaging. This report applies that distinction to one release and uses current records where later experiments supersede the historical overview.

2. Model and method

The three components are Unsloth's existing Q2_K diffusion GGUF, Sakura's modified Q4_K_M text encoder, and a BF16 VAE. The repository states that these files were copied byte-for-byte from pinned sources in provenance.json. The diffusion model is 7B-class; the separate text encoder is 8B-class. Q2_K denotes quantization, not a two-billion-parameter model. The documented inference starting settings are 8 steps, CFG 1.0, Euler and the upstream Turbo sigma schedule. [1, 2, 3, 4]

3. Evidence and results

The smoke-test record reports a generated 512 by 512 PNG for a neutral teapot prompt, seed 42, with stable-diffusion.cpp revision c6dfa9def12fd0cb36ac7149ba30a75792bedbf0. It records disk parameter storage and a managed GPU budget setting of 6 GiB. This is a basic execution record, not a measured peak-memory requirement, quality comparison or content-behavior evaluation. [1, 2, 3, 4]

4. Limitations and intended use

The modified encoder's label does not establish a measured content effect. Low-bit diffusion quantization can change outputs, but no numerical quality penalty is established here. Load all three components in a compatible runtime. The package is governed by the recorded Qwen Research License for noncommercial research/evaluation, with original component notices retained.

5. Reproducibility and conclusion

The source snapshot above fixes what was reviewed. The linked source files and source-manifest.json record the evidence used, including file hashes. Consult the repository's usage instructions and recorded dependencies for execution. No new model training, inference benchmark, or claimed state-of-the-art result was introduced in preparing this report. Evidence status: recorded smoke test.

References

[1] Lucasgitguy-co. README.md at reviewed revision.

[2] Lucasgitguy-co. provenance.json at reviewed revision.

[3] Lucasgitguy-co. smoke-test.json at reviewed revision.

[4] Lucasgitguy-co. docs/USAGE.md at reviewed revision.

[5] Lucasgitguy-co. Local Model Adaptation Across Vision Diagrams Geometry and Game Decisions. Historical portfolio report, 5 November 2026; its proposed v2 experiment is superseded by later release records.