

Zen Director: Text/Image-to-Video Generation

Zen Research Authors
Zen Research DAO, Zoo Labs Inc (501(c)(3))
San Francisco, CA • dev@hanzo.ai

September 2025

Abstract

Meta-study of zen-director based on CogVideoX-5B + Wan2.2-Ti2V, analyzing alternatives and selection rationale for the Zen AI ecosystem.

1 Introduction

This paper presents zen-director, evaluates alternatives in text/image-to-video generation, and justifies our selection of CogVideoX-5B + Wan2.2-Ti2V as the foundation.

2 Upstream Attribution

Based on **CogVideoX-5B + Wan2.2-Ti2V** [?]. We thank the original authors.
URL: <https://github.com/THUDM/CogVideo>

3 Selection Rationale

We selected CogVideoX-5B + Wan2.2-Ti2V for:

- State-of-the-art quality in text/image-to-video generation
- Open-source Apache 2.0/MIT licensing
- Production-grade stability and performance
- Active development and community support
- Extensibility for Zen ecosystem integration

4 Zen AI Ecosystem

Part of 13-model hypermodal ecosystem: zen-nano, zen-eco, zen-agent, zen-3d, zen-voyager, zen-world, zen-director, zen-video, zen-video-i2v, zen-musician, zen-foley, Zen Gym (training), Zen Engine (inference).

5 Conclusion

CogVideoX-5B + Wan2.2-Ti2V provides the best foundation for zen-director in the Zen ecosystem.