

Zen Director: Video Generation

Zen Research Authors
Zen Research DAO
Zoo Labs Inc (501(c)(3) Non-Profit)
San Francisco, California, USA
dev@hanzo.ai
+1 (913) 777-4443

September 2025

Abstract

5B text/image-to-video model.

1 Introduction

5B text/image-to-video model.

1.1 Key Features

- 5B parameter video model
- Up to 10 seconds, 24 FPS
- 1280x720 resolution
- Text and image conditioning

2 Technical Specifications

Parameter	Value
Parameters	5B
Duration	Up to 10 seconds
FPS	24
Resolution	1280x720
Input	Text, image

Table 1: Technical specifications

3 Zen AI Ecosystem

Part of the complete Zen AI hypermodal ecosystem:

Language: zen-nano-0.6b, zen-eco-4b, zen-agent-4b

3D & World: zen-3d, zen-voyager, zen-world

Video: zen-director-5b, zen-video, zen-video-i2v

Audio: zen-musician-7b, zen-foley

Infrastructure: Zen Gym (training), Zen Engine (inference)

4 Conclusion

5B text/image-to-video model.

Acknowledgments

Thanks to the open-source community and our upstream contributors.