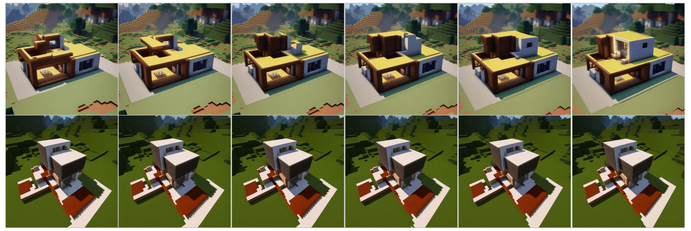
MagicTime: Time-lapse Video Generation Models as Metamorphic Simulators

Shenghai Yuan*, Jinfa Huang*, Yujun Shi, Yongqi Xu, Rujie Zhu, Bin Lin, Xinhua Chen, Li Yuan, Jiebo Luo

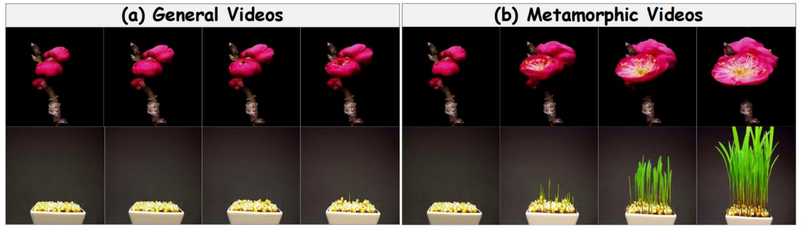
Motivation

Contributions:

- Encode more physical knowledge of the real world, thereby enabling the generation of metamorphic videos.
- Propose an automatic metamorphic video captioning annotation framework and curate a high-quality dataset named ChronoMagic.



What is Metamorphic Video?

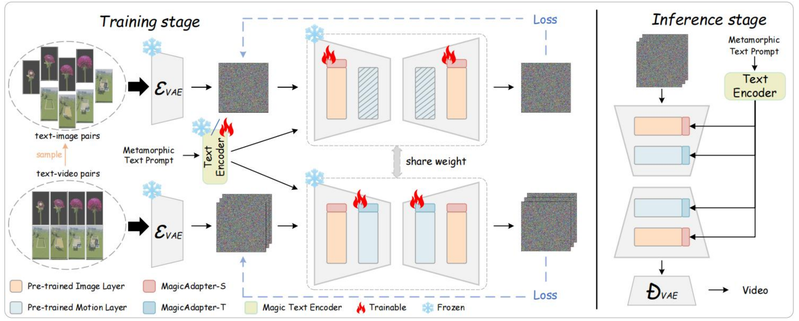


- Different from general video, the metamorphic video show strong physics, high persistence, and variation.

Dataset	Type	Strong Physics	High Persistence	Variation
Sky Time-Lapse [ACMMM'21] [8]	General	✗	✗	✗
Time-Lapse-D [CVPR'18] [9]	General	✗	✗	✗
ChronoMagic	Metamorphic	✓	✓	✓

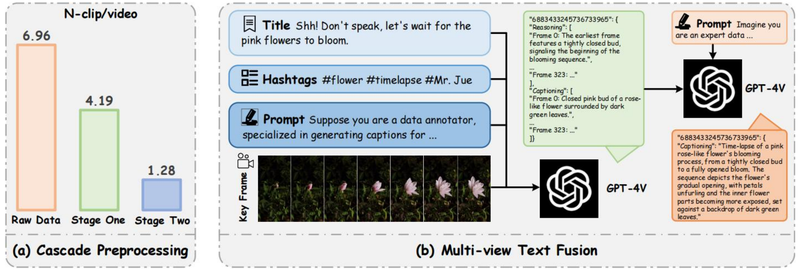
Framework

We decoupled the training of metamorphic and general motion, also adopted a dynamic frame extraction strategy more suitable for time-lapse videos to obtain training frames, and enhanced the text encoder's understanding of metamorphic prompts.



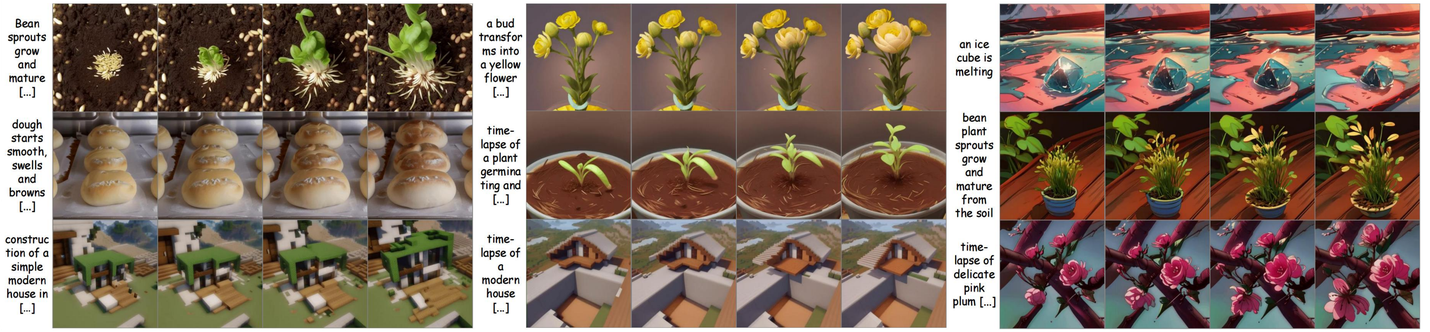
Annotation

Our video captioning annotation framework can automatically removes video transitions and can combine multi-view information to generate video annotations for metamorphic videos.



Visualization

- MagicTime can generalize well to personalized SD models to produce high-quality metamorphic videos.

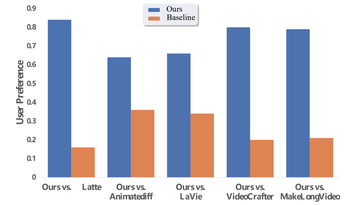


Results

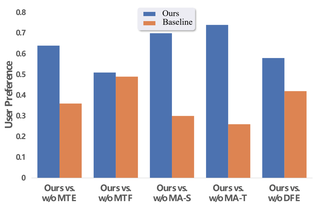
- Our model consistently outperforms the recently proposed state-of-the-art methods on metamorphic video generation benchmark.

Method	Venue	Backbone	FID↓	FVD↓	CLIPSIM↑
MakeLongVideo [58]	Github'23	U-Net	143.07	511.62	0.2852
ModelScopeT2V [5]	Arxiv'23	U-Net	61.49	500.43	0.3093
VideoCrafter [59]	CVPR'23	U-Net	**54.61**	473.69	0.3149
Zeroscope [60]	CVPR'23	U-Net	103.74	497.60	0.3079
LaVie [61]	Arxiv'23	U-Net	62.28	445.60	0.3090
T2V-zero [62]	ICCV'23	U-Net	58.58	488.85	0.3203
Latte [3]	Arxiv'24	DiT	59.10	445.85	0.2939
Animatediff [1]	ICLR'24	U-Net	63.83	451.35	0.3156
MagicTime	Ours	U-Net	58.12	**441.17**	**0.3272**

Metamorphic video generation performance on ChronoMagic.



Human Evaluation on ChronoMagic.



Human evaluation of different ablation settings

Personal

Paper

Demo

Code