Learning Video Representations without Natural Videos

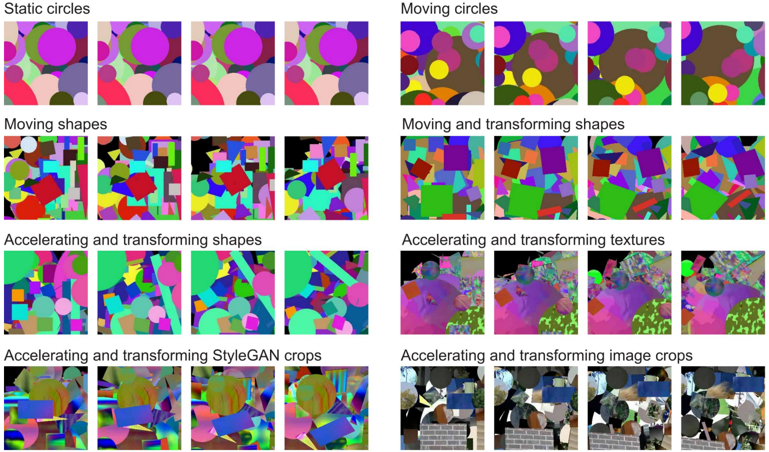
Xueyang Yu, Xinlei Chen, Yossi Gandelsman

ICCV HONOLULU HAWAII OCT 19-23, 2025

Motivation

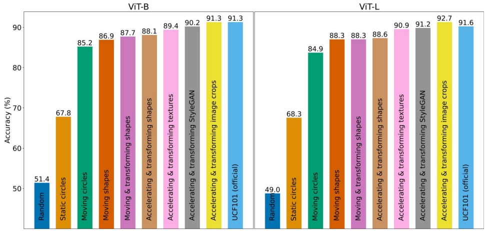
1. Self-supervised learning from videos require large amount of data and not yet achieve same success as images domain.
2. We question the the efficient utilization of current SSL models by curating synthetic datasets to match performance with real videos.
3. We further try to identify key properties in useful video data from large data corups.

Synthetic Data Progression



- We introduce a progression of synthetic video data from noise, textures and image crops and include motions.
- We pretrain VideoMAE with synthetic / real data and compare representation on downstream action recognition tasks.
- Using same # of pretrain videos, synthetic data and compete with real data and outperforms real data in OOD coruptions.
- We curate ~30 synthetic datasets and analyze how different attributes affect representation quality.

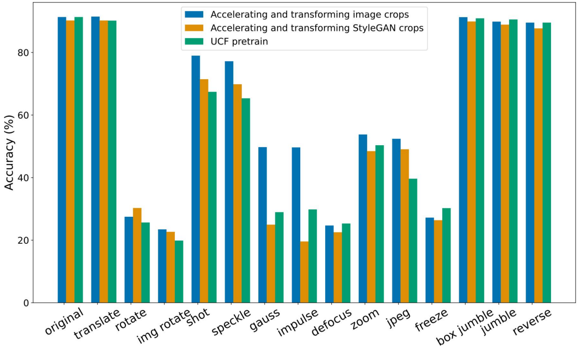
Downstream Task Evaluation



	HMDB51 fine-tune	UCF101 lin. prob	UCF101 fine-tune
Random initialization	18.2	8.9	51.4
Static circles	29.2	13.2	67.8
Moving circles	52.0	15.5	85.2
Moving shapes	56.1	20.4	86.9
Moving and transforming shapes	57.6	18.8	87.7
Acc. and transforming shapes	58.9	18.9	88.1
Acc. and transforming textures	62.4	20.9	89.4
Acc. and transforming StyleGAN crops	**64.1**	25.2	90.2
Acc. and transforming image crops	**64.1**	24.8	**91.3**
UCF101	63.0	**48.0**	**91.3**

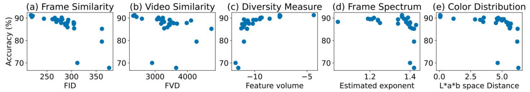
Synthetic data can compete or even outperform natural videos across different settings.

OOD Generalization



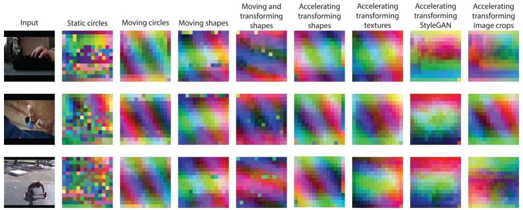
Synthetic data outperforms real videos in 11 out of 14 OOD benchmarks.

Dataset Attributes



1. More motions lead to stronger temporal information capture.
2. Higher diversity correlates with better representation.
3. Nature-style spectral and color statistics reduce distribution gaps.

Attention Visualization



Model Learned motion and semantic information increases with our data progression.

Paper

Project Page