

The Weizmann Institute of Science Faculty of Mathematics and Computer Science

Vision and AI

Lecture Hall - Room 1, Ziskind Building
on Thursday, May 14, 2026
at 12:15

Hila Chefer
BFL

will speak on

Self-Flow: Self-Supervised Flow Matching for Scalable Multi-Modal Synthesis

Abstract:

Despite the dominance of diffusion and flow-based models in visual and multi-modal generation, they remain paradoxically dependent on external encoders for strong semantic representations, leading to limited generalization and unexpected scaling behavior. This reliance suggests a fundamental limitation: standard local denoising provides little incentive for models to develop global semantic structure. To address this, we introduce Self-Flow, a self-supervised framework that unifies representation learning and generative modeling within a single pipeline. By utilizing Dual-Timestep Scheduling to create an information asymmetry across tokens, we force the model to infer corrupted inputs from cleaner context, naturally driving the emergence of strong representations alongside generative capabilities. Because Self-Flow operates without external models, it generalizes seamlessly across modalities and follows expected scaling laws, offering a path toward world models that are both perceptually grounded and semantically rich. Teaser video in this [Link](#).

Bio:

Dr. Hila Chefer is a Research Scientist at Black Forest Labs and an incoming Assistant Professor (Senior Lecturer) at Tel Aviv University. Her research focuses on architecture development and interpretability for visual foundation models, aiming to both understand and advance their generative capabilities. Hila earned her PhD from Tel Aviv University under the supervision of Prof. Lior Wolf, where she developed the de facto leading methods for Transformer interpretability and the Attend-and-Excite framework for text-to-image control. Her work in video generation includes the development of Lumiere during her time at Google and VideoJAM at Meta AI.

