CVPR 2025 Second Workshop on

Efficient and On-Device Generation (EDGE)

Time: June 12, 2025 Thursday (1-5 PM)

Location: 208 A

https://cvpr25-edge.github.io/

PixVerse

Speakers

Stefano Ermon

Stanford

Lu Liu

OpenAI

Ziwei Liu

Nanyang Technological University

Ishan Misra

GenAI at Meta

Sergey Tulyakov

Snap Inc.

Jiaming Song

Luma AI

Enze Xie

NVIDIA

Jun-Yan Zhu

Carnegie Mellon University

Lu Jiang

ByteDance

Organizers

Organizing Committee

Felix Juefei-Xu*

GenAI, Meta

Tingbo Hou*

Google DeepMind

Yang Zhao

Google DeepMind

Licheng Yu

GenAI, Meta

Zhisheng Xiao

Google DeepMind

Xiaoliang Dai

GenAI, Meta

Qifei Wang

Google DeepMind

Tao Xu

GenAI, Meta

Yanwu Xu

Google

Ali Thabet

GenAI, Meta

Qiang Liu

UT Austin

Xuan Ju

CUHK

Ruiqi Gao

Google DeepMind

Xi Yin

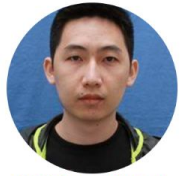
GenAI, Meta

Haolin Jia

Google

Xide Xia

GenAI, Meta



Peizhao Zhang

GenAI, Meta

Peter Vajda

GenAI, Meta

Schedule

Schedule

Each talk is assigned a 25 min slot (20 min talk + 4 min QA + 1 min buffer)

Time	Activity	Title
13:00 - 13:10	Opening remarks and award announcement	
13:10 - 13:35	Ziwei Liu, Nanyang Technological University	*"From Multimodal Generative Models to Dynamic World Modeling"*
13:35 - 14:00	Stefano Ermon, Stanford University	*"Accelerating Inference in Diffusion Models"*
14:00 - 14:25	Ishan Misra, GenAI at Meta	*"Scale Efficient Video Generation and Tokenization"*
14:25 - 14:50	Sergey Tulyakov, Snap Inc.	*"Sharpening the Edge: High Quality Image and Video Synthesis on Mobiles"*
14:50 - 14:55	**Break**	
14:55 - 15:20	Jiaming Song, Luma AI	*"Breaking the Algorithmic Ceiling in Pre-Training with a Inference-first Perspective"*
15:20 - 15:45	Jun-Yan Zhu, Carnegie Mellon University	*"Distilling Diffusion Models into Conditional GANs"*
15:45 - 16:10	Lu Jiang, ByteDance	*"Cost-Effective Training of Video Generation Foundation Model"*
16:10 - 16:35	Enze Xie, Nvidia	*"Building Image Generation model from scratch and Acceleration"*
16:35 - 17:00	Lu Liu, OpenAI	*"A Brief Introduction of 4o Image Generation"*

Papers

Accepted Long Papers (6 in total)

(1) Title: Geometric Consistency Refinement for Single Image Novel View Synthesis via Test-Time Adaptation of Diffusion Models
Authors: Josef Bengtson (Chalmers University of Technology), David Nilsson (Chalmers University of Technology), Fredrik Kahl (Chalmers University of Technology)
[Poster # 202]

(2) Title: AdaVid: Adaptive Video-Language Pretraining
Authors: Chaitanya Patel (Stanford University), Juan Carlos Niebles (Salesforce Research), Ehsan Adeli (Stanford University)
[Poster # 203]

(3) Title: LLMPi: Optimizing LLMs for High-Throughput on Raspberry Pi
Authors: Mahsa Ardakani (University of South Carolina), Jinendra Malekar (University of South Carolina), Ramtin Zand (University of South Carolina)
[Poster # 204]

Papers

Accepted Long Papers (6 in total)

(4) Title: Latent Patched Efficient Diffusion Model For High Resolution Image Synthesis
Authors: Weiyun Jiang (Rice University), Devendra Kumar Jangid (Samsung Research America), Seok-Jun Lee (Samsung Research America), Hamid Rahim Sheikh (Samsung Research America)
[Poster # 205]

(5) Title: ADAPTOR: Adaptive Token Reduction for Video Diffusion Transformers
Authors: Elia Peruzzo (University of Trento), Adil Karjauv (Qualcomm), Nicu Sebe (University of Trento), Amir Ghodrati (Qualcomm AI Research), Amir Habibian (Qualcomm)
[Poster # 206]

(6) Title: Scaling On-Device GPU Inference for Large Generative Models
Authors: Jiuqiang Tang (Google), Raman Sarokin (Google), Ekaterina Ignasheva (Meta), Grant Jensen (Google), Lin Chen (Google), Juhyun Lee (Google), Andrei Kulik (Google), Matthias Grundmann (Google)
[Poster # 207]

Papers

Accepted Short Papers (7 in total)

(1) Title: Random Conditioning for Diffusion Model Compression with Distillation
Authors: Dohyun Kim (Korea University), Sehwan Park (Korea University), Geonhee Han (Korea University), Seung Wook Kim (Nvidia), Paul Hongsuck Seo (Korea University)
[Poster # 208]

(2) Title: LatentLLM: Attention-Aware Joint Tensor Compression
Authors: Toshiaki Koike-Akino (Mitsubishi Electric Research Labs), Xiangyu Chen (Mitsubishi Electric Research Labs), Jing Liu (Mitsubishi Electric Research Labs), Ye Wang (Mitsubishi Electric Research Labs), Pu Perry Wang (Mitsubishi Electric Research Labs), Matthew Brand (Mitsubishi Electric Research Labs)
[Poster # 209]

(3) Title: MS-Temba: Multi-Scale Temporal Mamba for Efficient Temporal Action Detection
Authors: Arkaprava Sinha (University of North Carolina at Charlotte), Monish Soundar Raj (University of North Carolina at Charlotte), Pu Wang (University of North Carolina at Charlotte), Ahmed Helmy (University of North Carolina at Charlotte), Srijan Das (University of North Carolina at Charlotte)
[Poster # 210]

Papers

Accepted Short Papers (7 in total)

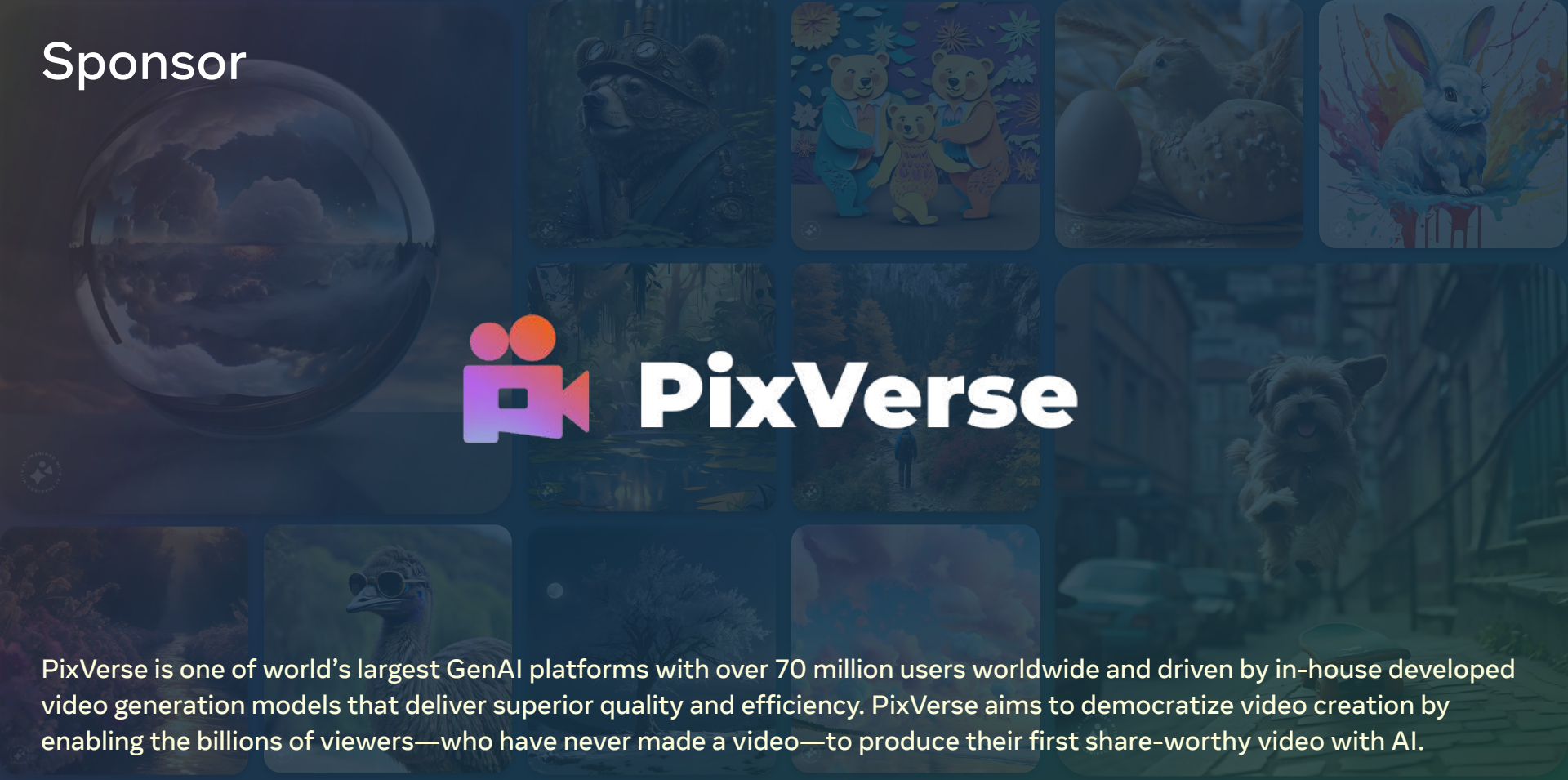
(4) Title: TuneComp: Joint Fine-Tuning and Compression for Large Foundation Models
Authors: Xiangyu Chen (Mitsubishi Electric Research Labs), Jing Liu (Mitsubishi Electric Research Labs), Ye Wang (Mitsubishi Electric Research Labs), Matthew Brand (Mitsubishi Electric Research Labs), Pu Perry Wang (Mitsubishi Electric Research Labs), Toshiaki Koike-Akino (Mitsubishi Electric Research Labs)
[Poster # 211]

(5) Title: Efficient Personalization of Quantized Diffusion Model without Backpropagation
Authors: Hoigi Seo (Seoul National University), Wongi Jeong (Seoul National University), Kyungryeol Lee (Seoul National University), Se Young Chun (Seoul National University)
[Poster # 212]

(6) Title: Atlas: Multi-Scale Attention Improves Long Context Image Modeling
Authors: Kumar Krishna Agrawal (University of California, Berkeley), Long Lian (University of California, Berkeley), Longchao Liu (University of California, Berkeley), Natalia Harguindeguy (University of California, Berkeley), Boyi Li (Nvidia Research), Alexander G Bick (Vanderbilt University), Maggie Chung (University of California, San Francisco), Trevor Darrell (University of California, Berkeley), Adam Yala (University of California, Berkeley)
[Poster # 213]

(7) Title: LoRA.rar: Learning to Merge LoRAs via Hypernetworks for Subject-Style Conditioned Image Generation
Authors: Donald Shenaj (Samsung R&D Institute UK (SRUK), University of Padova), Ondrej Bohdal (Samsung R&D Institute UK (SRUK)), Mete Ozay (Samsung R&D Institute UK (SRUK)), Pietro Zanuttigh (University of Padova), Umberto Michieli (Samsung R&D Institute UK (SRUK))
[Poster # 214]

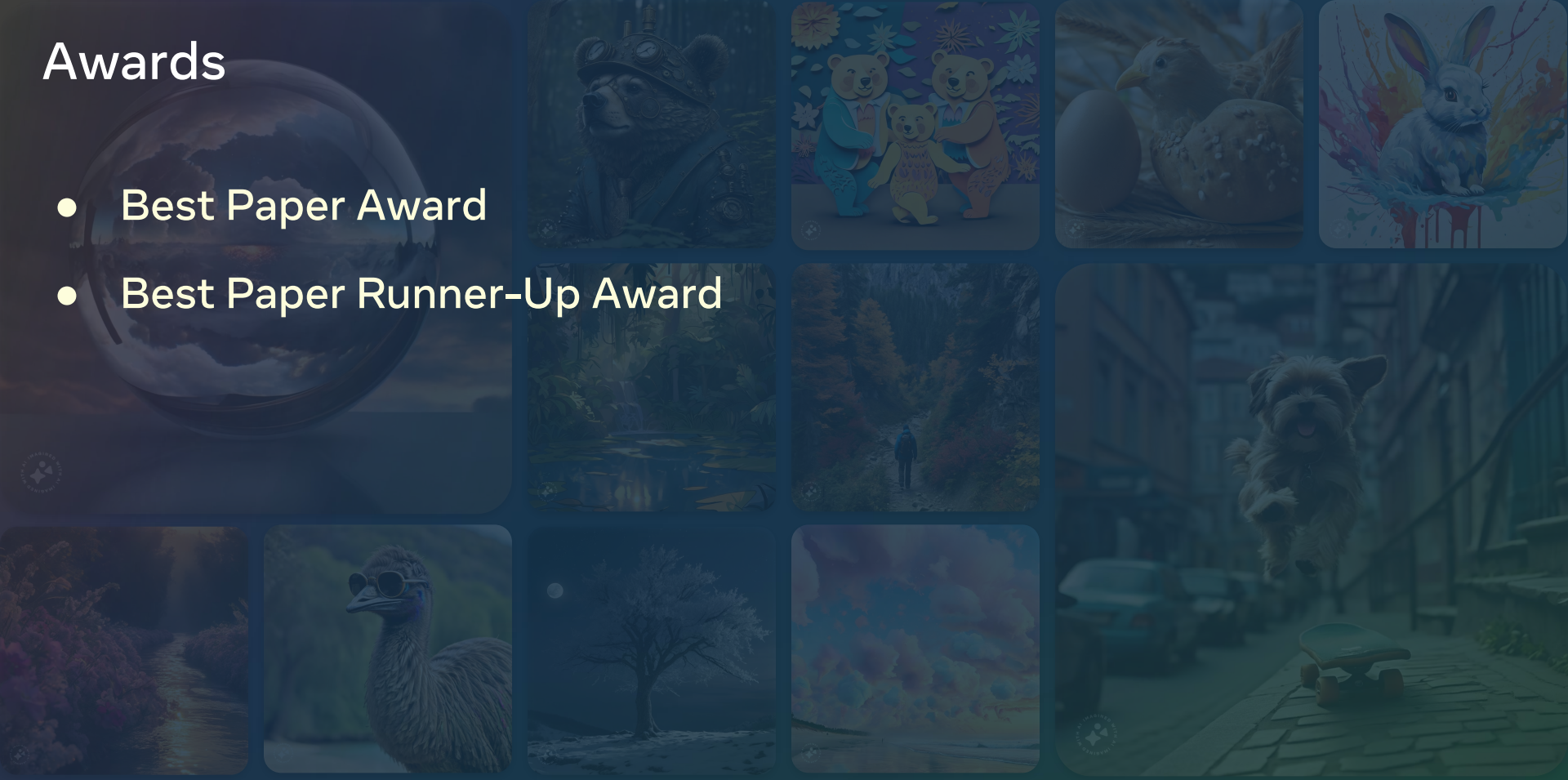
Sponsor



PixVerse is one of world's largest GenAI platforms with over 70 million users worldwide and driven by in-house developed video generation models that deliver superior quality and efficiency. PixVerse aims to democratize video creation by enabling the billions of viewers—who have never made a video—to produce their first share-worthy video with AI.

Awards

- Best Paper Award
- Best Paper Runner-Up Award




CVPR 2025 EDGE Workshop
Best Paper

CVPR 2025 EDGE WORKSHOP BEST PAPER

Presented to

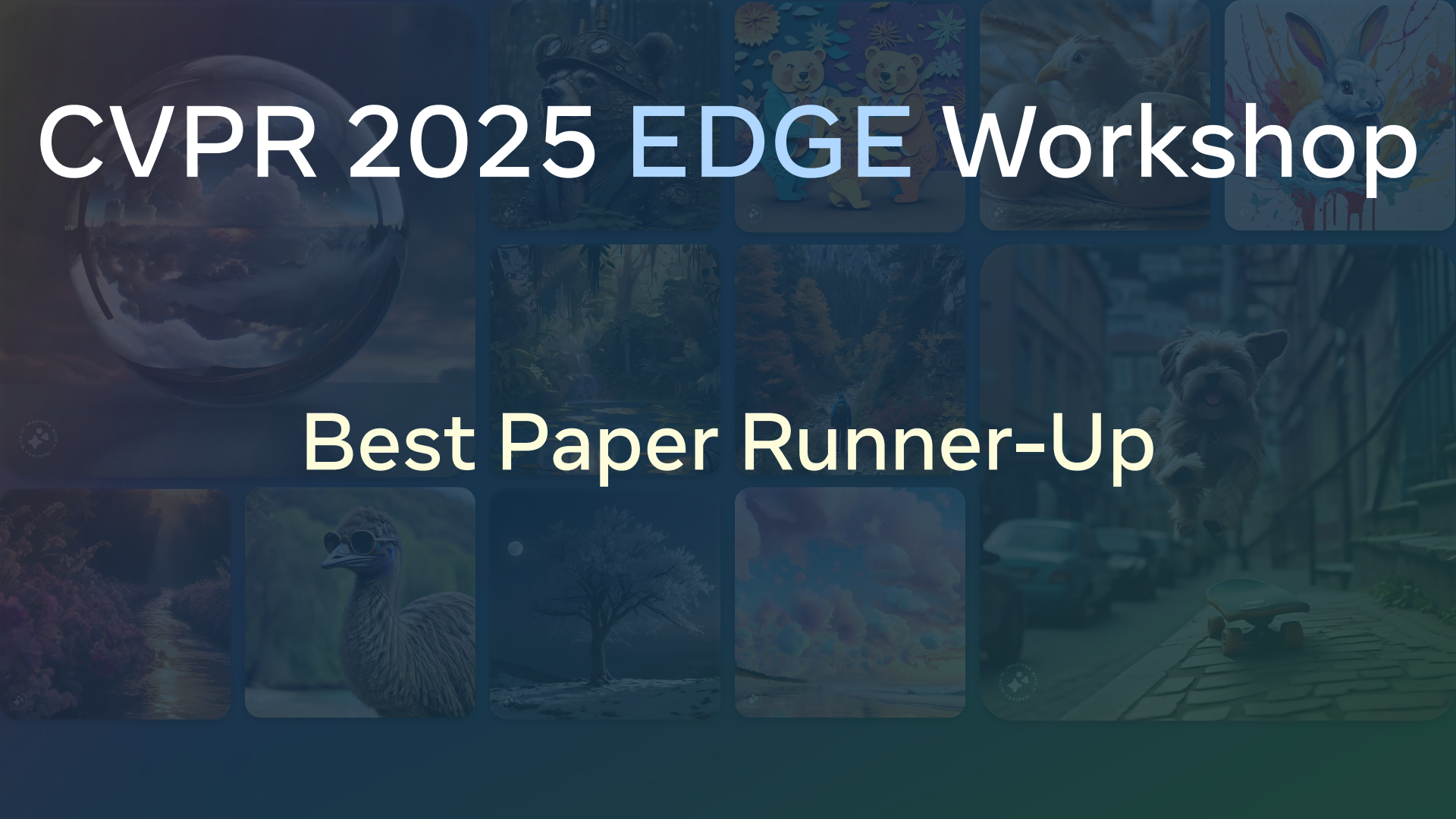
Chaitanya Patel, Juan Carlos Niebles, Ehsan Adeli

For the Paper

AdaVid: Adaptive Video-Language Pretraining

Felix Juefei-Xu
CVPR 2025 EDGE
Organizer

Tingbo Hou
CVPR 2025 EDGE
Organizer


CVPR 2025 EDGE Workshop
Best Paper Runner-Up

CVPR 2025 EDGE WORKSHOP
BEST PAPER RUNNER-UP

Presented to

Jiuqiang Tang, Raman Sarokin, Ekaterina Ignasheva, Grant Jensen, Lin Chen, Juhyun Lee, Andrei Kulik, Matthias Grundmann

For the Paper

Scaling On-Device GPU Inference for Large Generative Models

Felix Juefei-Xu
CVPR 2025 EDGE
Organizer

Tingbo Hou
CVPR 2025 EDGE
Organizer

Enjoy the Workshop!

Schedule

Each talk is assigned a 25 min slot (20 min talk + 4 min QA + 1 min buffer)

Time	Activity	Title
13:00 - 13:10	Opening remarks and award announcement	
13:10 - 13:35	Ziwei Liu, Nanyang Technological University	*"From Multimodal Generative Models to Dynamic World Modeling"*
13:35 - 14:00	Stefano Ermon, Stanford University	*"Accelerating Inference in Diffusion Models"*
14:00 - 14:25	Ishan Misra, GenAI at Meta	*"Scale Efficient Video Generation and Tokenization"*
14:25 - 14:50	Sergey Tulyakov, Snap Inc.	*"Sharpening the Edge: High Quality Image and Video Synthesis on Mobiles"*
14:50 - 14:55	**Break**	
14:55 - 15:20	Jiaming Song, Luma AI	*"Breaking the Algorithmic Ceiling in Pre-Training with a Inference-first Perspective"*
15:20 - 15:45	Jun-Yan Zhu, Carnegie Mellon University	*"Distilling Diffusion Models into Conditional GANs"*
15:45 - 16:10	Lu Jiang, ByteDance	*"Cost-Effective Training of Video Generation Foundation Model"*
16:10 - 16:35	Enze Xie, Nvidia	*"Building Image Generation model from scratch and Acceleration"*
16:35 - 17:00	Lu Liu, OpenAI	*"A Brief Introduction of 4o Image Generation"*


PixVerse