

## 01 Motivation & Problem

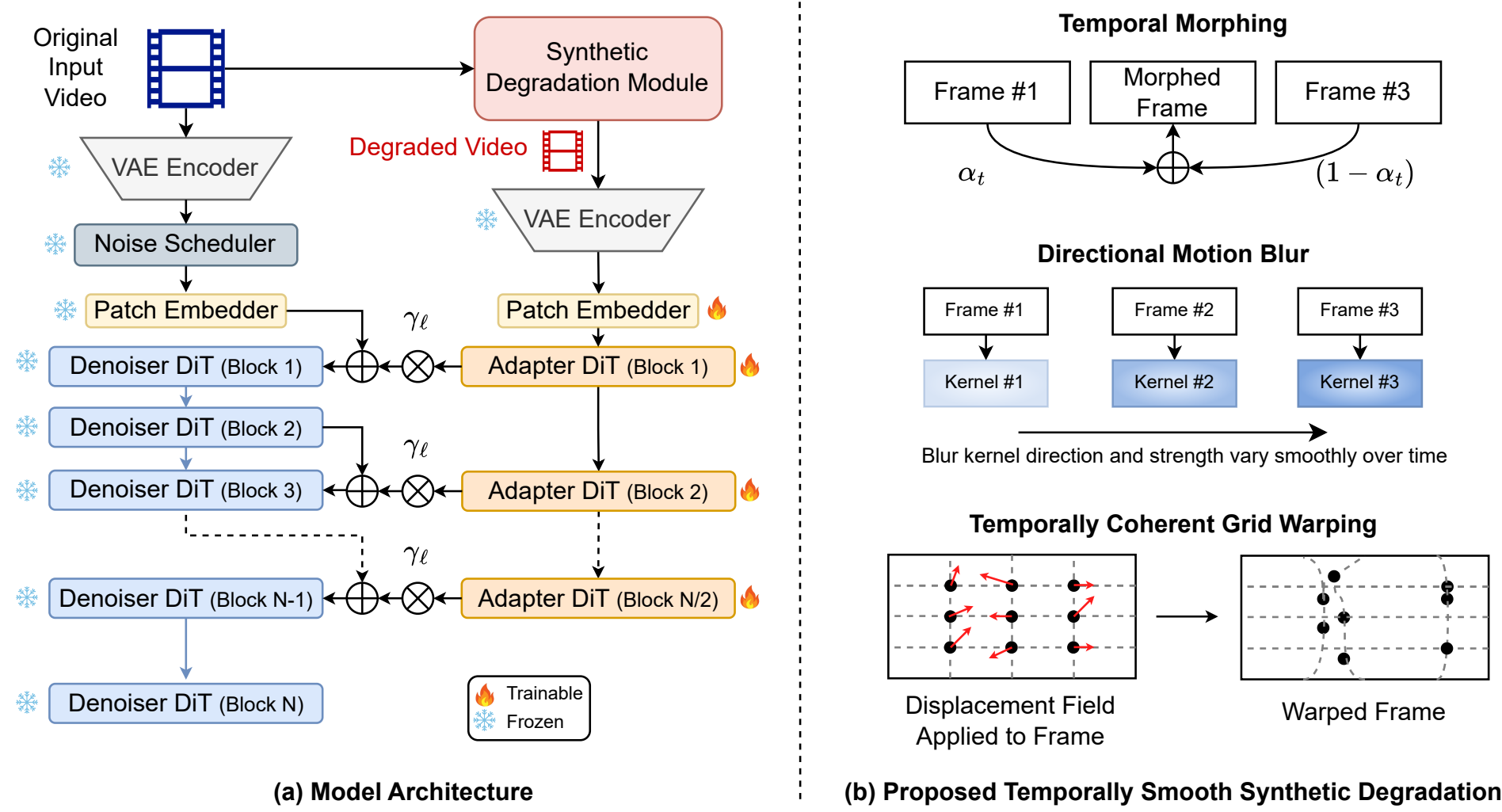
Modern video generators (Veo3, Sora2, Ray3, ...) still struggle with **fine detail** such as faces, hands, text, and motion.



## 02 Contributions

- Novel framework to repair structure and motion, beyond denoising and upscaling across both AI-generated and real video.
- A temporally-coherent degradation module for training a diffusion model that synthesizes realistic structural and motion failures
- A new AIGC54 benchmark for artifact-laden generated video, with multi-aspect FIQA, semantic, and perceptual scoring.
- State-of-the-art on severe-artifact video and competitive on standard VR/VSR, at a practical ~13 FPS for 720p on a single A100.

## 03 Method – The Pipeline

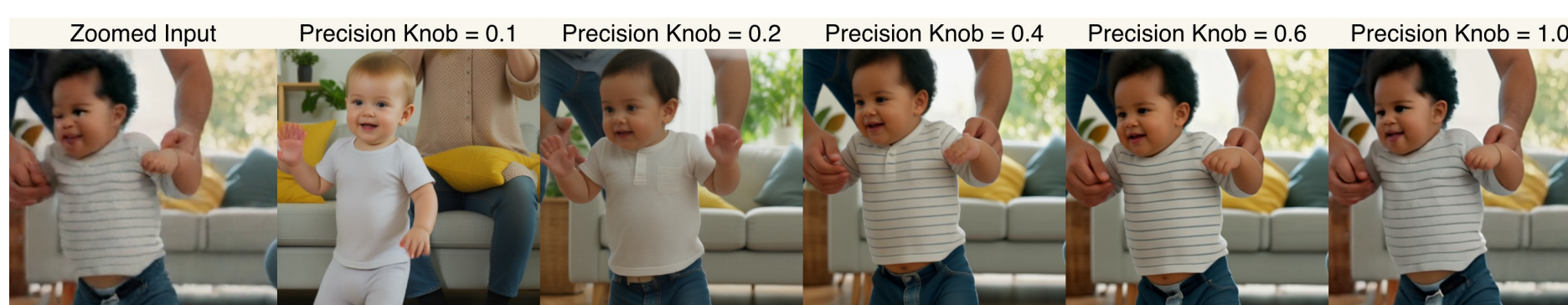


### ★ Key Novelty: Synthetic degradation module

Generates temporally-coherent structural and motion artifacts, training the model to actively correct them.

## 04 The Precision Knob ( $\gamma$ )

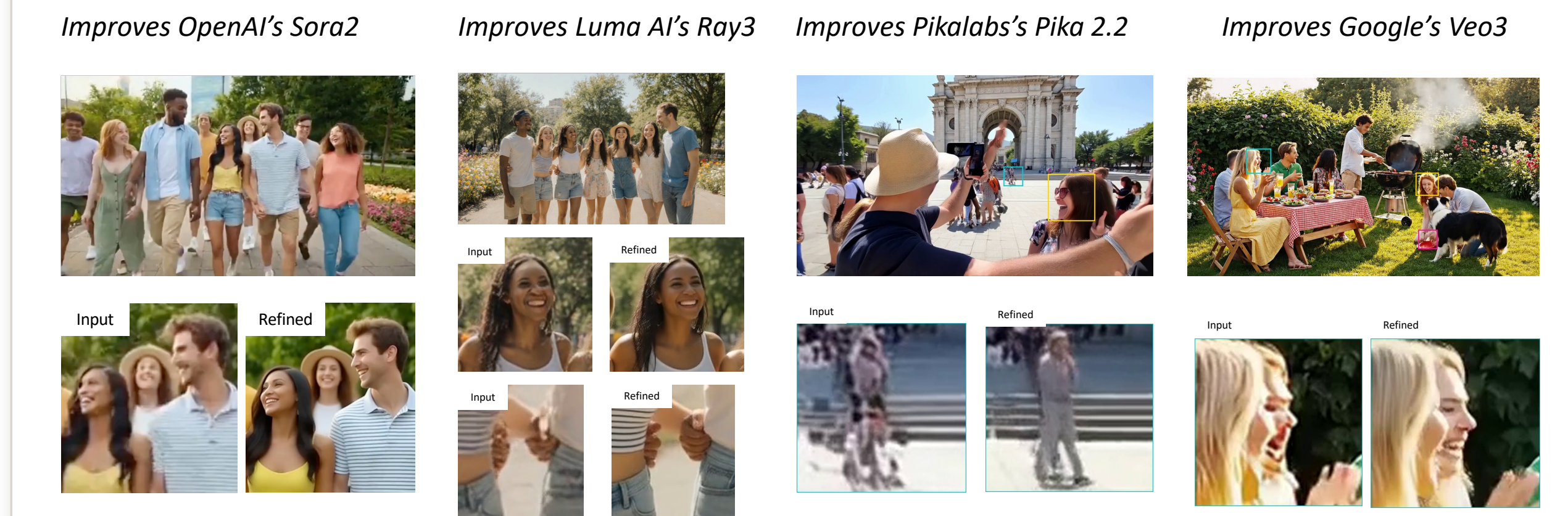
High precision preserves input details; low precision enables stronger corrective synthesis



## 05 Quantitative Benchmarks

- **AIGC benchmark (ours):** clips from frontier video generators
  - Metrics: FIQA (face/structural quality), VBench, and V-LLM as judge.
- **Standard VSR:** UDM10, SPMCS, REDS30
  - Metrics: measured by PSNR, SSIM, LPIPS.
- **Result:** SOTA on artifact-heavy AIGC video; competitive on standard VSR.

## 06 Results – Improves all frontier Video Generation models



## 07 Additional Applications



## 08 References

1. Zhuang, Junhao, et al. "FlashVSR: Towards Real-time Diffusion-Based Streaming Video Super Resolution." CVPR, 2026
2. Wang, Jianyi, et al. "Seedvr2: One-step video restoration via diffusion adversarial post-training." arXiv, 2025
3. Zhang, Yuehan, and Angela Yao. "Realviformer: Investigating attention for real-world video super-resolution." ECCV, 2024.



PROJECT PAGE