

# Diffusion Models Are Real Time Game Engines

Dani Valevski\*, Yaniv Leviathan\*, Moab Arar\*, Shlomi Fruchter\*



## Abstract

Game engine powered by a diffusion model that simulates DOOM in Real-Time

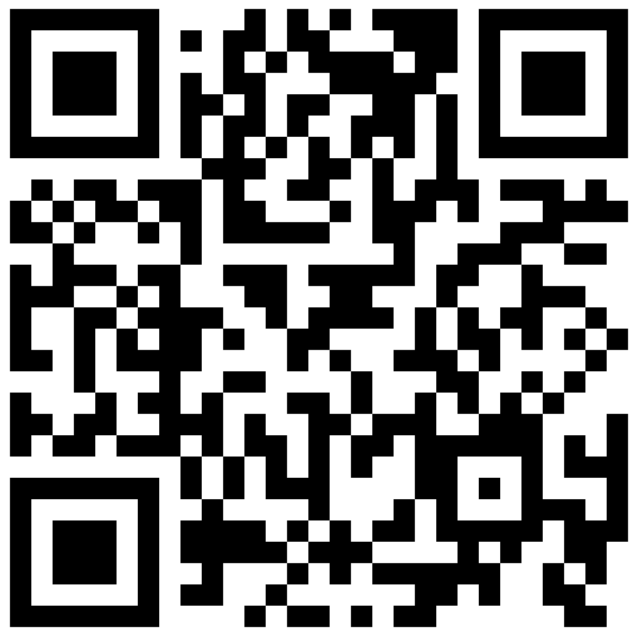
## Key Features

20FPS Long Simulations

High Quality Generation 29.4 PSNR

Game Logic Understanding

## Gameplay Recording



## Pipeline Overview

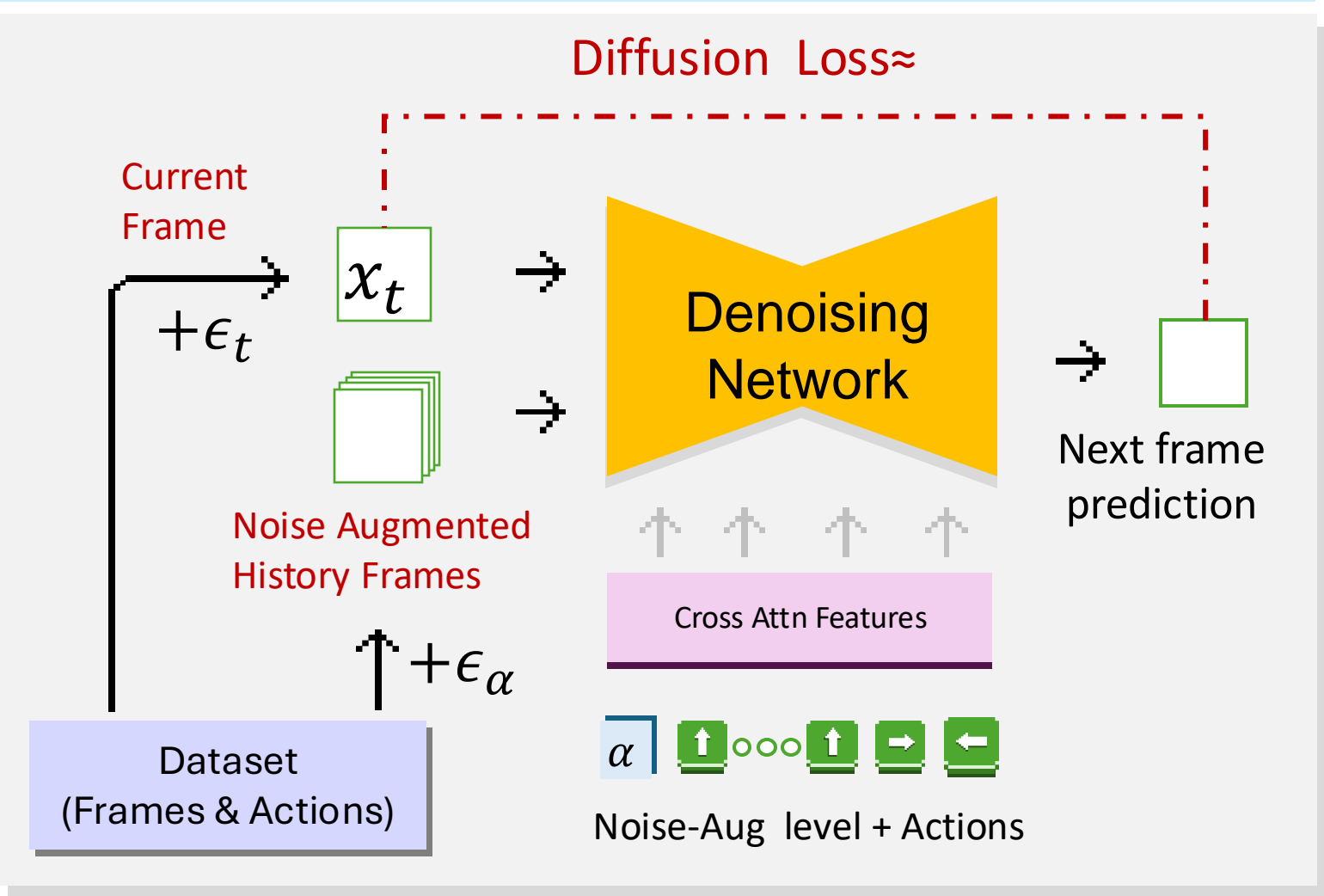
### Data Collection Via RL Agent



Conditioned on Prev frames & actions

VAE Decoder Finetuning for fine-grained details

### Finetuning SD V1.4



Noise augmentation for autoregressive Stabilization

4 Denoising Steps for real-time generation

## VAE Decoder Finetuning



Before



After

## Noise Augmentation



Before

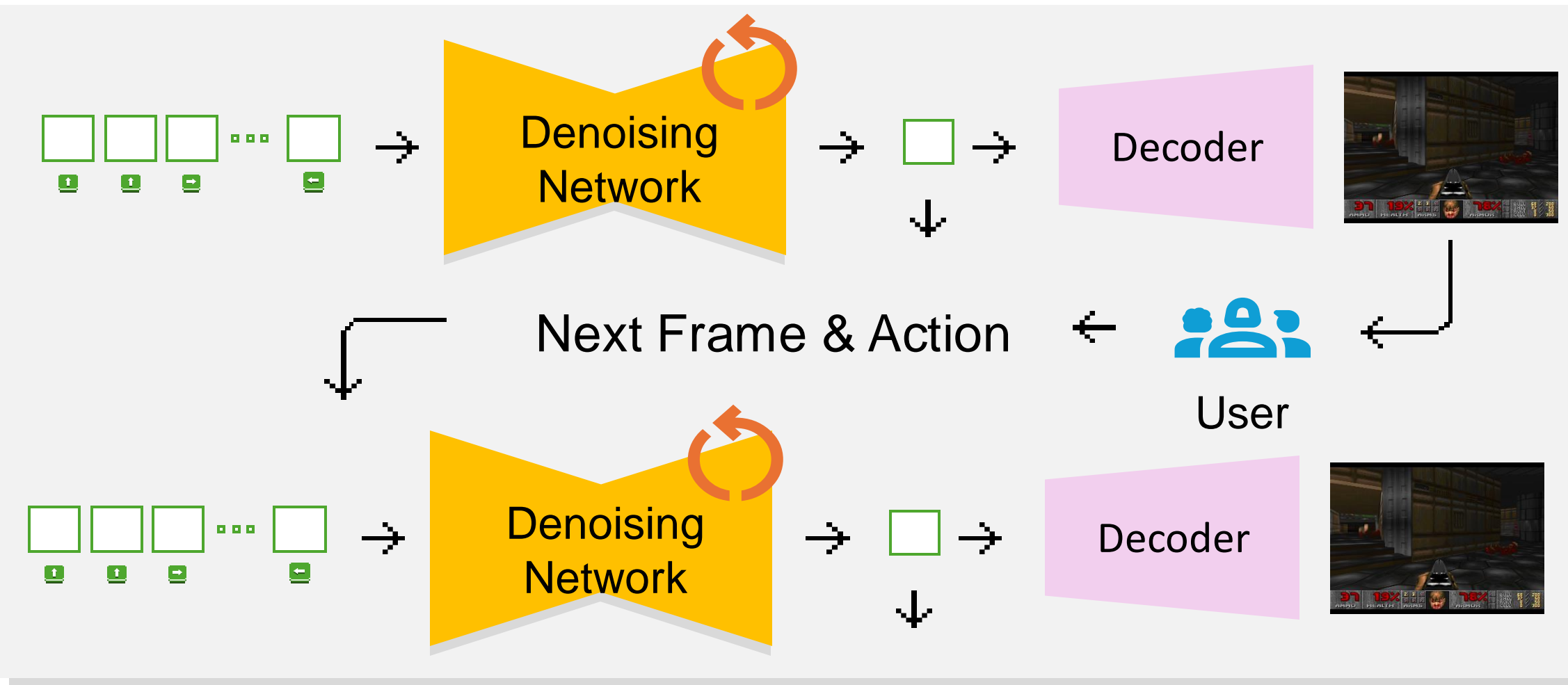


After

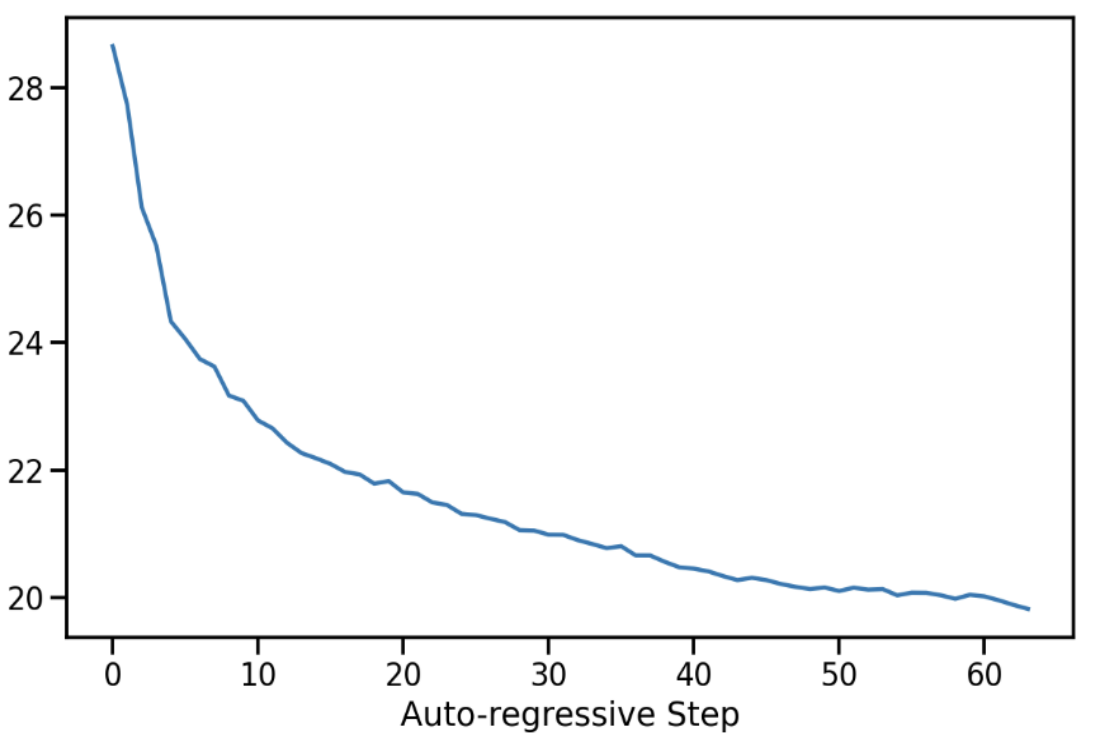
## Real-time Gameplay Simulated by GameNGen



## Inference



## PSNR vs Autoregressive Step



## Denoising Steps vs Quality

Steps	PSNR ↑	LPIPS ↓
D	31.10 ± 0.098	0.208 ± 0.002
1	25.47 ± 0.098	0.255 ± 0.002
2	31.91 ± 0.104	0.205 ± 0.002
4	32.58 ± 0.108	0.198 ± 0.002
8	32.55 ± 0.110	0.196 ± 0.002
16	32.44 ± 0.110	0.196 ± 0.002
32	32.32 ± 0.110	0.196 ± 0.002
64	32.19 ± 0.110	0.197 ± 0.002