Scene Integrated Generation for Neural Radiance Fields

Uses ControlNets for Depth-Image Conditioning

Align your Latents: High-Resolution Video Synthesis with Latent Diffusion Models (Nvidia)

Photorealistic Video Generation with Diffusion Models

Google, Stanford, 2023