Netflix VOID: Removing Objects from Video Without Breaking the Physics

Kabui, Charles

2026-04-04

Netflix researchers released VOID (Video Object and Interaction Deletion), a framework that removes objects from video while fixing the physical consequences of their absence. Current video removal tools can erase objects and clean up shadows or reflections, but they fail when the removed object physically interacts with other things: a ball knocking over a cup, a person holding a guitar, a weight pressing down on a pillow. VOID handles this with a two-stage

approach. First, a vision-language model identifies which parts of the scene were causally affected by the removed object. Then a video diffusion model, fine-tuned on CogVideoX, generates what the scene would have looked like if the object never existed. Training data comes from Kubric and HUMOTO, physics simulators where you can actually remove objects and re-simulate to get ground-truth counterfactual video. An optional second pass uses flow-warped noise to fix object morphing artifacts, a common failure mode of smaller diffusion models.

A VFX artist removing a product placement or an unwanted extra from footage currently has to manually repaint every frame where the removed object affected something else. VOID automates that counterfactual reasoning. The code, model weights, and a live demo are all public. It requires a GPU with 40GB+ VRAM (an A100 or better), so this is a professional tool, not a consumer one yet.

Video editing models are being pushed toward genuine world simulation. Removing an object is no longer just an inpainting problem. It is a physics reasoning problem, and the tools are starting to reflect that.

Sources:

- VOID: Video Object and Interaction Deletion (arXiv)
- Netflix/void-model (GitHub)
- VOID Project Page
- VOID Demo (HuggingFace Spaces)
- VOID Model Weights (HuggingFace)
