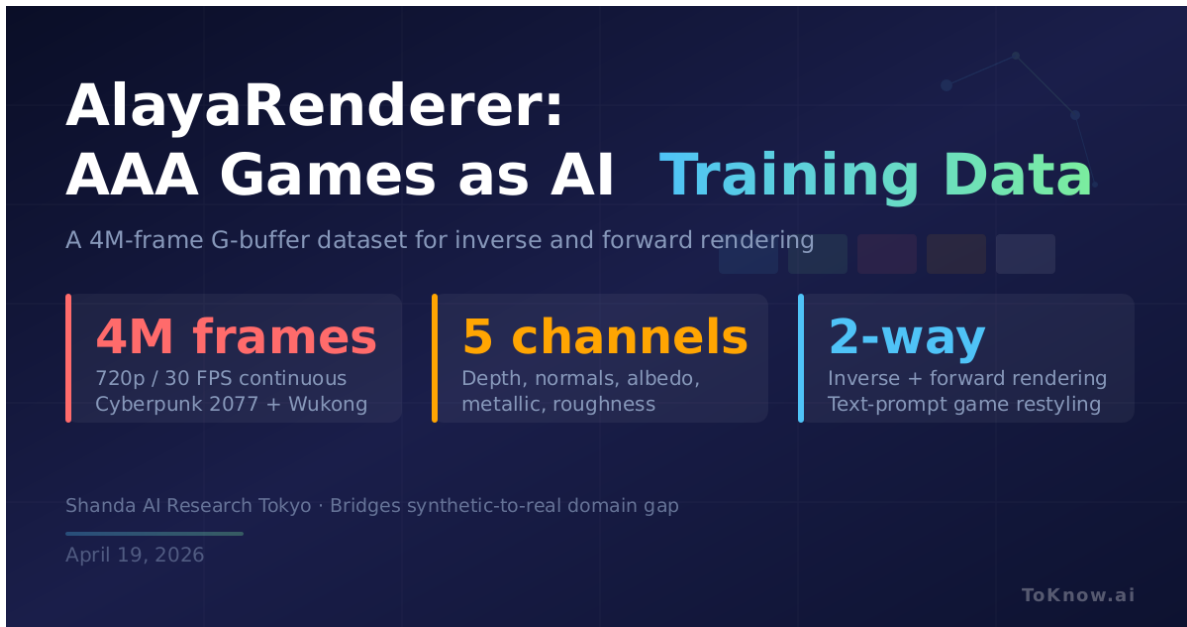


AlayaRenderer: 4 Million AAA Game Frames That Teach AI to See Light, Materials, and Geometry

Kabui, Charles

2026-04-19

[Read at ToKnow.ai](#)



AlayaRenderer:
AAA Games as AI Training Data

A 4M-frame G-buffer dataset for inverse and forward rendering

- 4M frames**
720p / 30 FPS continuous
Cyberpunk 2077 + Wukong
- 5 channels**
Depth, normals, albedo,
metallic, roughness
- 2-way**
Inverse + forward rendering
Text-prompt game restyling

Shanda AI Research Tokyo · Bridges synthetic-to-real domain gap
April 19, 2026
ToKnow.ai

Shanda AI Research Tokyo released [Generative World Renderer \(AlayaRenderer\)](#), a toolkit and dataset that pulls 4 million continuous frames at 720p and 30 FPS from two AAA games, [Cyberpunk 2077](#) and [Black Myth: Wukong](#), each frame paired with five [G-buffer](#) channels:

depth, surface normals, albedo, metallic, and roughness. These channels are the raw geometric and material layers a game engine computes before lighting, and they are what an AI model needs to learn how images decompose back into shape, material, and light. A custom dual-screen capture method extracts these channels from commercial games that don't normally expose them. Models fine-tuned on the dataset deliver cleaner [inverse rendering](#) (turning a photo back into geometry, material, and lighting) on real-world scenes, and the same data drives a forward renderer that lets users restyle game environments through text prompts.

The realism gap has been the bottleneck for AI rendering. Lab-built synthetic datasets use simple objects and clean lighting, so models trained on them break on real photos with weather, motion blur, and complex materials. AAA games already render at near-photoreal quality with full ground-truth geometry, which makes them a far better training source than anything academia can build from scratch. The team also proposes a [vision-language model](#) evaluation that scores consistency without needing ground truth, and reports it tracks human judgment closely. For a film studio or game team, this means relighting footage, swapping weather, or restyling a scene from a single video clip, with no manual 3D modelling.

Read More: [WildWorld: 108 Million Frames from Monster Hunter for Training World Models](#) covers Shanda AI's earlier dataset that used the same games-as-data approach for action-conditioned world models.

Sources:

- [Generative World Renderer \(project page\)](#)
- [arXiv:2604.02329 — Generative World Renderer](#)
- [GitHub: ShandaAI/AlayaRenderer](#)
- [HuggingFace Daily Papers](#)
- [Hugging Face Playground: Game Editing](#)

*Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. **Read more:** [/terms-of-service](#)*