Google Earth’s AI deepfake tool only lasted one day

Google has shut down Google Earth feature it launched Thursday that allowed users to edit satellite images with text prompts using AI. The tool essentially let users create AI deepfakes of the real world using text prompts; *Digital Digging*’s Henk van Ess, for example, intentionally generated images adding things like refugees near the Mexican border and a bomb crater by a hospital in Gaza. On Thursday, Google’s initial response to van Ess noted that images generated with Nano Banana 2 in Google Earth included a digital watermark and that the company prevented “image creation on harmful topics.”

One day later, Google’s new statement says that “We know that people uniquely trust Google Earth for a reliable view of the world. We’ve seen geospatial professionals using this feature for a range of useful purposes, however we’ve also seen people sharing screenshots of generated imagery that appear to violate our policies. So we’re rolling back this feature in Google Earth while we work on implementing stronger guardrails. It’s important to note that generated images didn’t appear in the main Google Earth experience for others to see and were watermarked as AI generated.”

van Ess pointed out that he was able to fool Hive’s AI detection tool with a video generated in Google Earth, showing how the generated media could spread despite the watermarks, and said in his testing, “Nothing was refused, nothing was softened, and nothing suggested I try a different prompt.”

https://www.theverge.com/tech/973943/google-earth-ai-image-generation-deepfake-tool