Artists are taking things into their own hands to protect their work from generative AI

July 18 2024, by Sarah Parvini

Credit: Google DeepMind from Pexels

The oil painting depicts a woman standing on a podium, her arm aloft as she grasps a laurel crown in her hand. A scarlet cloak drapes across her chest as she stares at the viewer. To the naked eye, the painting looks like a normal piece in an online portfolio. But the version of the painting uploaded online belies a hidden defense system—a tool called Glaze that masks the artist's style and cloaks the art from use by generative AI.

As image-generating AI continues to evolve, artists have increasingly fought against what they see as an existential threat to their craft on multiple fronts: through lawsuits, in public statements calling for regulations and now, with programs aimed at protecting their art from being scraped and emulated without their permission.

Created by researchers from the University of Chicago, Glaze and a second program, Nightshade, essentially poison the well of art uploaded online in an attempt to scramble what AI sees. Where Glaze subtly changes an image so AI perceives it as a different art style, Nightshade is a more "offensive" tool that attempts to confuse an AI training model about what is in an image. The idea, university researchers said, is to provide a technical solution to stop the "malicious" use of AI models while also protecting creators.

The team working on Glaze and Nightshade noted that their tools serve as a safeguard in a space that lacks regulation rather than as a comprehensive solution. Where lawsuits and government regulations might force tech giants like Microsoft or OpenAI to change how they operate, smaller AI players outside of the United States might not follow suit, researchers said. In those instances, the AI scrambling tools will still be useful.

"Regulations, strikes are very important and arguably launch a much more profound, long lasting effect on this entire landscape," said Shawn Shan, the lead student working on Glaze and Nightshade. "But I think we see Glaze, and specifically Nightshade, as more leverage."

For Karla Ortiz, the San Francisco-based artist behind the oil painting and the first person to publicly use Glaze, the intersection of art and AI "all boils down to consent."

"It's just deeply unfair to work your whole life to train, to learn to be able to do the things that we do to find our own voice as artists, and then to have someone take that voice, make a mimic of it, and say, 'Oh, actually, this is ours,'" she said. "Artists need a way to be able to exist online."

Ortiz is one of three artists seeking to protect their copyrights and careers by suing makers of AI tools that can generate new imagery on command. The suit against Stability AI, the London-based maker of text-to-image generator Stable Diffusion, alleges that the AI image-generators violate the rights of millions of artists by ingesting huge troves of digital images and then producing derivative works that compete against the originals.

Ortiz, a concept artist and illustrator in the entertainment industry who has worked on movies including "Rogue One: A Star Wars Story" and "Doctor Strange," said technology like Glaze is key to protecting artists.

"It's this disgusting cycle of tech saying we own what is yours, but you don't get to have a say in how you use your work," said Ortiz. "What you post online, that's ours. And we're also going to compete in your markets. And we're gunning for your job, which is essentially what's happening."

Experts say anti-AI tools do provide some protection by making it harder for people trying to use AI to mimic an artist's style, but they don't eradicate the problem. As AI models evolve, they will likely become harder to attack or throw off.

"When AI becomes stronger and stronger, these anti-AI tools will become weaker and weaker," said Jinghui Chen, an assistant professor at Pennsylvania State University who co-authored a study on the effectiveness of tools like Glaze. "But I recognize that as a first step."

It's important, Chen said, to raise questions about the efficacy of these types of tools in order to improve them.

Shan, the University of Chicago researcher, agreed that the anti-AI tools are "far from future-proof."

"But this is the case for most of the security mechanisms that we see in the digital age," he said. "Firewalls, they are not perfect. There are many ways to bypass them. But most people still use firewalls to stop a good amount of these attacks, these types of problems. So we see Glaze or Nightshade similarity."

Artists shouldn't "blindly trust" that all their problems will be solved by those tools, he added.

Renato Roldan, an artist who uses Glaze before uploading his work to his portfolio, said that he's hesitant to update as much as he used to because "we are super exposed now" with AI. He worries about how generated art will shift the paradigm on how art is consumed and created, he said, likening an image created by AI to a diluted version of art created by a person.

"If you do a photocopy of something, it starts to deteriorate every time you copy it," said Roldan, a former video game art director who now works in narrative and storyboarding.

One of the biggest consequences of unregulated AI, he said, is that it has made it harder for new artists who are just finishing school to break into the field because they have to compete with generated art.

"They'd usually find a learning curve that was a little bit hard," he said. "But now they are finding that curve is not a curve. It's just a wall."

Citation: Artists are taking things into their own hands to protect their work from generative AI (2024, July 18) retrieved 19 July 2024 from https://techxplore.com/news/2024-07-artists-generative-ai.html