

Engineers look to an old source to empower the future of computer vision

July 7 2023, by Adam Hadhazy



Princeton researchers have developed an open-source software system that generates an infinite number of photorealistic scenes of the natural world, an advance that could improve the training of autonomous cars and other robots. Image courtesy the researchers. Credit: Princeton University

Artificial intelligence seems perfect for creating massive sets of images needed to train autonomous cars and other machines to see their environment, but current generative AI systems have shortcomings that can limit their use. Now, engineers at Princeton have developed a software system to overcome those limits and quickly create image sets to prepare machines for nearly any visual setting.

The new system, called [Infinigen](#), relies on mathematics to create natural looking objects and environments in three dimensions. Infinigen is a procedural generator, which in [computer science](#) denotes a program that creates content based on automated, human-designed algorithms rather than labor-intensive manual data entry or the [neural networks](#) that power modern AI. In this way, the new program generates myriad 3D objects using only randomized mathematical rules.

Infinigen is "a dynamic program for building unlimited, diverse, and realistic natural scenes," said Jia Deng, an associate professor of computer science at Princeton and senior author of a new study that details the [software system](#). The paper was presented at the [CVPR 2023](#) conference.

Infinigen's mathematical approach allows it to create labeled visual data, which is needed to train [computer vision systems](#), including those deployed on home robots and autonomous cars. Because Infinigen generates every image programmatically—it creates a 3D world first, populates it with objects, and places a camera to take a picture—Infinigen can automatically provide detailed labels about each image including the category and location of each object.

The images with automatic labels can then be used to train a robot to recognize and locate objects given only an image as input. Such labeled visual data would not be possible with existing AI image generators, according to Deng, because those programs generate images using a deep neural network that does not allow the extraction of labels.

In addition, Infinigen's users have fine-grained control of the system's settings, such as the precise lighting and viewing angle, and can fine-tune the system to make images more useful as training data.

Besides generating virtual worlds populated by digital objects with

natural shapes, sizes, textures and colors, Infinigen's capabilities extend to synthetic representations of natural phenomena including fire, clouds, rain and snow.

"We expect that Infinigen will prove to be a useful resource not just for creating training data for computer vision, but also for augmented and virtual reality, game development, film-making, 3D printing, and content generation in general," Deng said.

To build Infinigen, the Princeton researchers started with Blender, a free-to-use, open-source graphic system of prebuilt software tools that dates to the 1990s. In keeping with the spirit of Blender, the Princeton researchers have released Infinigen's code under a GPL-compatible license, meaning anyone can freely use it.

By vastly expanding the menu of 3D-rendered objects and landscapes, another key advantage of Infinigen is that it can boost machines' ability to perform 3D reconstructions, from just 2D pixels, of the complex spaces they will operate within. While moving away from real-world images to synthetic images to develop cars and robots that will move in the real world might seem counterintuitive, real image datasets have key limitations, Deng said.

For starters, the computers that guide robots and smart cars do not perceive images and other visual objects like humans do. An image that looks three-dimensional to a human is just a two-dimensional collection of pixels to a computer. To allow robots to perceive an image in 3D, the image needs to include an instruction called a "3D ground truth." This is difficult to do with existing 2D images, but easy for a system like Infinigen.

"Synthetic datasets of 3D images have shown great initial promise," said Deng, "and we developed Infinigen to further deliver on this promise."

For Infinigen, the Princeton researchers designed subprograms, dubbed generators, that specialize in producing single distinct types of digital objects—for instance, "fish" or "mountains." Users can work with the subprograms to tailor a range of parameters including size, texture, color and reflectivity.

"Users can tweak the parameters to create as much realness or unrealness as they desire for their particular task," said Deng. "The expansiveness can help ensure that machines are being broadly trained to handle and navigate the full spectrum of encounterable environments."

The researchers hope that Infinigen will become a collaborative tool, allowing users to add more features as it develops.

"A goal is for Infinigen coverage to become so good that the project becomes the go-to place for computer vision [training data](#), whatever the task is," said Deng. "We want Infinigen to become a collaborative, community-driven effort that provides a useful tool for a lot of users."

More information: Report: [Infinite Photorealistic Worlds Using Procedural Generation](#)

Provided by Princeton University

Citation: Engineers look to an old source to empower the future of computer vision (2023, July 7) retrieved 11 September 2024 from <https://techxplore.com/news/2023-07-source-empower-future-vision.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.