# Copy and paste: New AI tool helps computers interpret the world

December 13 2023, by Greg Hardesty



Overall pipeline of physically plausible object insertion for monocular 3D object detection: Our approach copies external 3D objects (e.g., from Objaverse [Deitke et al., 2022]) and pastes them into indoor scene datasets (e.g., SUN RGB-D [Song et al., 2015]) in a physically plausible manner. The augmented indoor scene dataset, enriched with inserted 3D objects, is then used to train monocular 3D object detection models, resulting in significant performance improvements. Credit: Yunhao Ge et al, 3D Copy-Paste: Physically Plausible Object Insertion for Monocular 3-D Detection (2023)

Copy and paste: It's a simple concept. You define some text or image on your computer, copy it, and paste it where you want it. Now, think of that new leather sofa you crave. Popular augmented reality (AR) apps allow you to cut and paste an image of the sofa into a photo of your living room to see if you like it before buying.

A team of researchers at USC Viterbi's Thomas Lord Department of Computer Science has now developed a similar technique to copy virtual 3D objects and paste them into real indoor scenes. This creates an overall natural and realistic image in terms of spatial relationships, object orientations and lighting.

What's more, the technique—called 3D Copy-Paste—can teach computers how to recognize the virtual 3D object in a multitude of different settings without having to rely on the tedious and expensive process of having a human feed the computer with reams of data.

"This is about training machine-learning systems how to recognize 3D objects in indoor scenes with a method that significantly improves existing 3D object models and achieves state-of-the-art performance," said computer science Professor Laurent Itti.

One of Itti's doctoral students, Yunhao "Andy" Ge, is presenting a research paper, 3D Copy-Paste: Physically Plausible Object Insertion for Monocular 3D Detection, at the 37th Conference on Neural Information Processing Systems (NeurIPS 2023) in New Orleans, Dec. 11-16.

"This is the first paper to show that we can insert photo-realistic 3D objects into a real-world indoor scene and create enough data to train an AI model to scale up recognition of such objects on its own," Ge said.

Itti and Ge collaborated on the project with Assistant Professor of Computer Science Jiajun Wu and his fourth-year Ph.D. student at Stanford University, Hong-Xing "Koven" Yu, as well as four computer scientists with Bosch Research North America: Cheng Zhao, Yuliang Guo, Xinyu Huang, and Liu Ren.

## 'Profound' implications

The 3D Copy-Paste tool is what's known in the AI world as a generative data augmentation technique, in which algorithms are taught to produce coherent and meaningful content that closely resembles human-created output by learning from patterns, trends, and relationships.

3D Copy-Paste could have "profound" implications for both the computer graphics and computer vision fields, Itti and Ge said.

Take, for example, autonomous driving technology.

An image of a cow is most associated with pastures and other bucolic settings.

If you want to teach an AI in a self-driving car to avoid hitting a cow in front of your moving vehicle, the AI initially might get confused—a cow normally isn't found in the middle of a road. You would have to feed it an image of a cow in front of a car for it to recognize the object quickly.

But the 3D Copy-Paste tool allows a [computer](#) to recognize an object in an endless variety of environments without having to be frontloaded with a ton of images. And it can create new images that don't exist in the real world—say, a cow walking on the moon—that blend in seamlessly with a photo of an indoor environment and appear to be physically plausible.

"You don't need any human to do manual labeling," Ge explained, "because when this virtual 3D object is inserted into a real indoor scene, it automatically generates labels for the AI to understand."

Added Itti, "This tool can generate millions of combinations of an image of an object, which allows the AI model to be trained that much better because of the high-quality data this tool creates."

The key is making the inserted object physically plausible, which means

it won't "collide" with existing objects and will have the correct lighting. 3D Copy-Paste first identifies physically feasible locations and poses for the inserted objects to prevent collisions with the existing room layout. Subsequently, it estimates spatially varying illumination for the insertion location, enabling the immersive blending of the virtual objects into the original scene with plausible appearances and shadows.

## Virtual additions

In short, 3D Copy-Paste can improve how computers see and interpret things in 3D space.

"As AR technology becomes more widespread and used in various applications," Ge said, "the techniques we've developed can help enhance the user experience and make virtual objects blend seamlessly into our real world."

Another application of 3D Copy-Paste could be in the digitization of industrial workflows.

As industrial enterprises shift toward digitizing their workflows and creating digital twins of real-world assets, the ability to insert realistic 3D objects into these digital representations becomes crucial, Itti and Ge said.

The 3D Copy-Paste method, they said, could ensure that any virtual additions to these digital twins, such as new equipment or structures, are done in a physically accurate and visually coherent manner.

"Our findings highlight the potential of 3D data augmentation in improving the performance of 3D perception tasks, opening up new avenues for research and practical applications," Ge said.

**More information:** Yunhao Ge et al, [3D Copy-Paste: Physically Plausible Object Insertion for Monocular 3-D Detection](https://techxplore.com/news/2023-12-ai-tool-world.html) (2023).

Provided by University of Southern California