

Revolutionizing image generation through AI: Turning text into images

September 2 2022



Image generated from the text "Happy vegetables waiting for supper.". Credit: Ludwig Maximilian University of Munich

Creating images from text in seconds—and doing so with a conventional graphics card and without supercomputers? As fanciful as it may sound, this is made possible by the new Stable Diffusion AI model. The underlying algorithm was developed by the Machine Vision & Learning Group led by Prof. Björn Ommer (LMU Munich).

"Even for laypeople not blessed with artistic talent and without special computing know-how and [computer hardware](#), the new model is an effective tool that enables computers to generate images on command. As such, the model removes a barrier to [ordinary people](#) expressing their creativity," says Ommer. But there are benefits for seasoned artists as well, who can use Stable Diffusion to quickly convert new ideas into a variety of graphic drafts. The researchers are convinced that such AI-based tools will be able to expand the possibilities of creative image generation with paintbrush and Photoshop as fundamentally as computer-based word processing revolutionized writing with pens and typewriters.

In their project, the LMU scientists had the support of the start-up Stability.Ai, on whose servers the AI model was trained. "This additional computing power and the extra training examples turned our AI model into one of the most powerful image synthesis algorithms," says the computer scientist.

The essence of billions of training images

A special aspect of the approach is that for all the power of the trained model, it is nonetheless so compact that it runs on a conventional graphics card and does not require a supercomputer such as was

formerly the case for image synthesis. To this end, the [artificial intelligence](#) distills the essence of billions of training images into an AI model of just a few gigabytes.

"Once such AI has really understood what constitutes a car or what characteristics are typical for an [artistic style](#), it will have apprehended precisely these salient features and should ideally be able to create further examples, just as the students in an old master's workshop can produce work in the same style," explains Ommer. In pursuit of the LMU scientists' goal of getting computers to learn how to see—that is to say, to understand the contents of images—this is another big step forward, which further advances basic research in machine learning and computer vision.

The [trained model](#) was recently released free of charge under the "CreativeML Open RAIL-M" license in order to facilitate further research and application of this technology more widely. "We are excited to see what will be built with the [current models](#) as well as to see what further works will be coming out of open, collaborative research efforts," says doctoral researcher Robin Rombach.

More information: Robin Rombach et al, [High-Resolution Image Synthesis with Latent Diffusion Models](#), *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2022)

Provided by Ludwig Maximilian University of Munich

Citation: Revolutionizing image generation through AI: Turning text into images (2022, September 2) retrieved 26 April 2024 from <https://techxplore.com/news/2022-09-revolutionizing-image-ai-text-images.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.