

Researchers explore photographic images synthesized from semantic layouts

August 19 2017, by Nancy Owano



AI will serve to develop a network control system that not only detects and reacts to problems but can also predict and avoid them. Credit: CC0 Public Domain

How far can we go in achieving fictional scenes just by using real photos? More precisely, what can we do with deep learning in rendering video games? Those questions are the focus of research work by Qifeng Chen and Vladlen Koltun.

Their work attracted interest this month by *New Scientist* and other sites, exploring their approach. "It's paint by numbers for creating dreamy worlds," said *Engadget*.

Indeed, a video's notes said this was a paint by numbers approach to create a new image and it starts with a labeled layout. Sections are labeled as trees or cars, for example. The center might be labeled [road](#).

Luke Dormehl in *Digital Trends* described their work as having "[artificial intelligence](#) that can create photorealistic Google Street View-style images of fake street [scenes](#)."

The key operative is artificial intelligence. Matt Reynolds in *New Scientist* said the AI from Qifeng Chen of Stanford and Intel "works from rough layouts that tell it what should be in each part of the image." AI uses this layout as a guide to generate a completely new image.

The AI was trained on 3000 images of German streets, Reynolds said.

Digital Trends discussed their use of a "cascaded refinement network, a type of neural network designed to synthesize HD images with a consistent structure. Like a regular [neural network](#), a [cascaded refinement network](#) features multiple layers, which it uses to generate features one layer at a time."

With some human help it can build slightly blurry made-up scenes, said Roberto Baldwin, senior editor, *Engadget*. "To create an image a human needs to tell the AI system what goes where. Put a car here, put a building there, place a tree right there. It's paint by numbers and the system [generates](#) a wholly unique scene based on that input."

So fundamentally, Reynolds said, you are getting a fictional street that "was generated by an imaginative [neural](#) network, stitching together its

memories of real streets it was trained on."

"Chen's AI isn't quite good enough to create photorealistic scenes just yet," said Baldwin. However, it could be used to create video game and VR worlds "where not everything needs to look perfect in the near future."

Its creators think it could eventually be used for creating photorealistic video game worlds.

What's next? The researchers detailed their work in "Photographic Image Synthesis with Cascaded Refinement Networks," by Chen and Koltun, which is on arXiv.

They described their approach as synthesizing photographic images conditioned on semantic layouts. Using an "input layout," they achieved a rendering engine. The result is a corresponding photographic image.

The authors pointed to what was special about their work. "We show that photographic images can be [synthesized](#) from semantic layouts by a single feedforward [network](#) with appropriate structure, trained end-to-end with a direct regression objective."

They stated in their paper that "Exciting work remains to be done to achieve perfect photorealism. If such level of realism is ever achieved, which we believe to be possible, alternative routes for image synthesis in computer graphics will open up."

More information: Photographic Image Synthesis with Cascaded Refinement Networks, arxiv.org/abs/1707.09405

Citation: Researchers explore photographic images synthesized from semantic layouts (2017, August 19) retrieved 19 April 2024 from <https://techxplore.com/news/2017-08-explore-images-semantic-layouts.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.