

PCMag editors select and review products [independently](#). If you buy through affiliate links, we may earn commissions, which help support our [testing](#).

PCMag UK > Artificial Intelligence

AI Image Generator Can Also Produce Music (With Otherworldly Results)

The Riffusion program has been trained to generate spectrograms of any music you'd like, which can then be converted into audio clips.

by [Michael Kan](#)

Dec 15, 2022



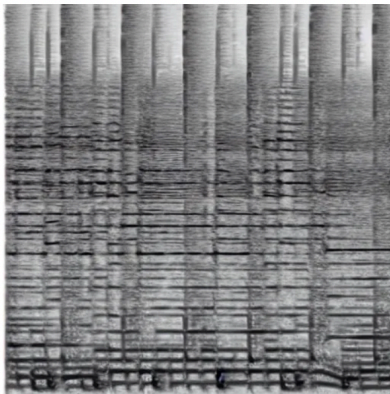
[AI image generators](#) can apparently do more than pump out professional art. A pair of bandmates are using the same algorithms to create music.

The project, called Riffusion, is using the AI text-to-image generation to produce spectrograms—or visual representations of music—which can then be converted into audio clips.

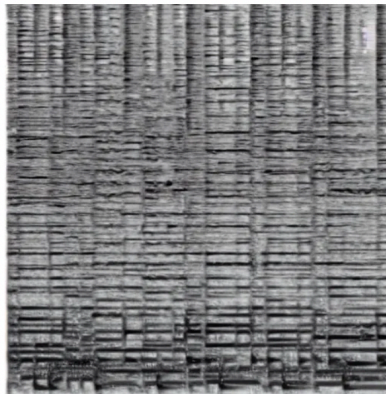
The AI trains computer algorithms to recognize pictures of objects and places, and uses them to replicate similar images. Programs including DALL-E 2, Midjourney, and Stable Diffusion are so adept at image generation they can try to visualize anything you want through various art styles—based on a mere text prompt.

The image generation inspired software developer Seth Forsgren and roboticist Hayk Martiros to see if the same AI programs could apply to the audio realm. “Hayk and I play in a little band together, and we started the project simply because we love music,” Forsgren tells PCMag. “Seeing the awesome results of Stable Diffusion for image generation, we asked ourselves what it would look like to use a diffusion approach to create music.”

rock and roll electric guitar solo

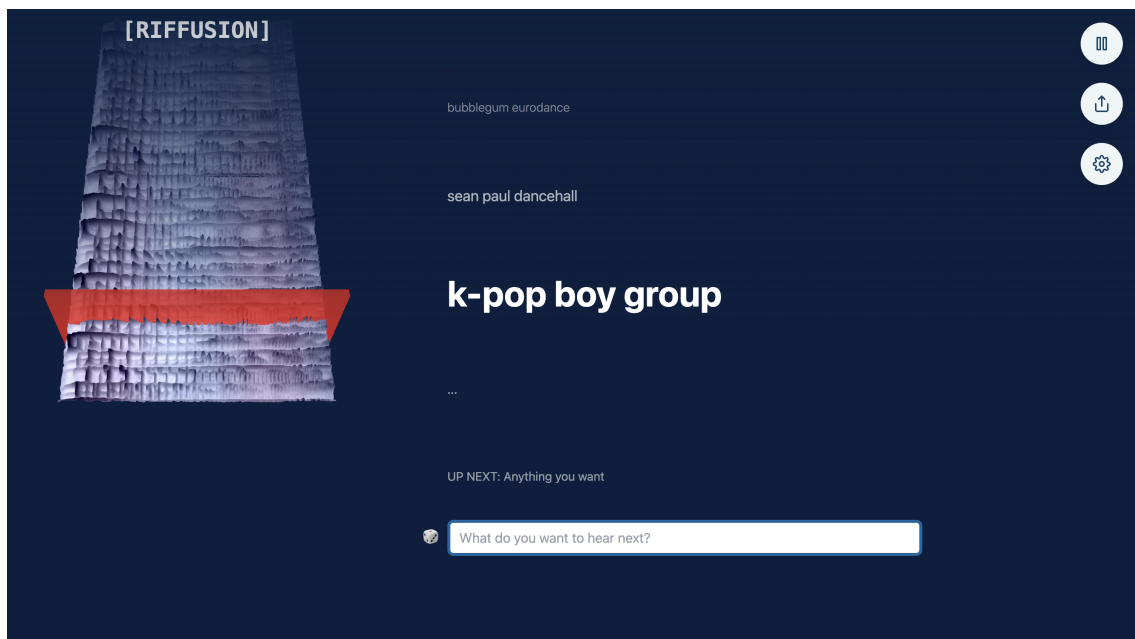


acoustic folk fiddle solo



To find out, the two trained the open-source Stable Diffusion on images of spectrograms paired with text. The program was then able to produce spectrograms of music based on a given prompt.

“At first we didn’t know if it would even be possible for the Stable Diffusion model architecture to create a spectrogram image with enough fidelity to convert into audio, but it turns out it can do that and more,” Forsgren says. “At every step along the way we’ve been more and more impressed by what is possible, and one idea leads to the next.”



Forsgren and Martiros published their [results](#) on the Riffusion website, which is meant to be a hobby project. But most importantly, visitors to the site can [plug in their own text prompts](#), and Riffusion will produce a spectrogram of their request, which will play on the site as an audio clip.

In addition, the program can create new variations of the spectrogram as you to listen. Here's an example of Riffusion trying to create an "Arabic gospel."

Your browser does not support the audio element.

The results are surprisingly good. We enjoyed this jazzy snippet, which was produced using the prompt: "funk bassline with a jazzy saxophone solo."

Your browser does not support the audio element.

Riffusion can also try to replicate songs, including K-Pop or an "[Eminem style anger rap](#)," without the lyrics. Instead, the tunes will feature melodic human-sounding gibberish that still matches the overall tone of the song.

For example, below is a "Fantasy ballad, female voice" that morphs into a "teen boy pop star" tune. To us, the resulting song sounds both human and alien at the same time.

Your browser does not support the audio element.

Forsgren says the lyrics from the program can sound "a bit otherworldly." Another limitation "is that the model is not designed to understand higher level song structure yet—like it doesn't try to repeat choruses or anything like that. You could imagine building an abstract model on top of this one to do that."

So the technology can't quite replace human-created music. But the project shows AI image algorithms may have plenty of unlocked potential for other purposes, including possibly offering music creators some inspiration. Forsgren and Martiros made Riffusion public on Thursday, and many users are already checking out their project.

"We are seeing a huge amount of traffic to the website —it's been a whirlwind effort to keep enough GPUs running to service all the requests, but we're having a fun day," Forsgren says. Hence, the Riffusion site may struggle to churn out your request due to the traffic load. However, the two also posted the code for their project on [GitHub](#).

MORE INSIDE PCMAG.COM

- [How to Pre-Order Google's New Pixel 9 Phones and Get the Best Deals](#)
- [3 Things Companies Must Get Right to Keep Hybrid Workers Happy](#)
- [Tonight's the Night: Perseid Meteor Shower Reaches Its Fiery Peak](#)
- [Figure's Humanoid Robot Gets a Sleek Upgrade](#)
- [Meta to Pay \\$1.4 Billion to Settle Texas Facial-Recognition Lawsuit](#)

About Michael Kan



I've been with PCMag since October 2017, covering a wide range of topics, including consumer electronics, cybersecurity, social media, networking, and gaming. Prior to working at PCMag, I was a foreign correspondent in Beijing for over five years, covering the tech scene in Asia.

More From Michael Kan