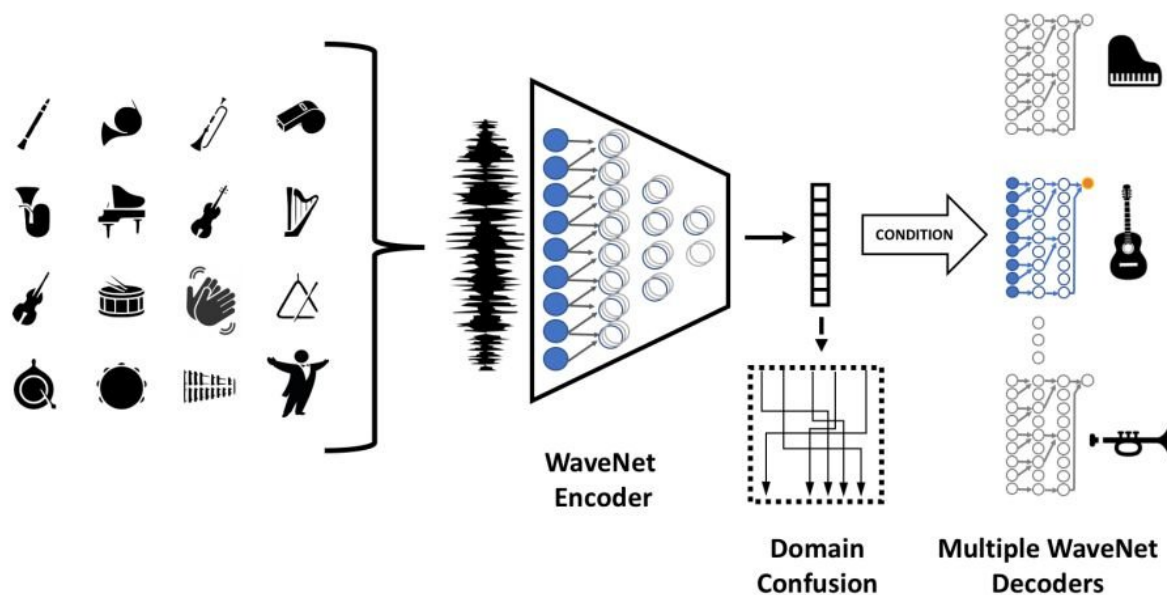# Translating instruments, styles, genres at Facebook Artificial Intelligence Research

May 28 2018, by Nancy Owano

The architecture of our network. The confusion block (dashed line) is employed only during training. Credit: arXiv:1805.07848 [cs.SD]

Is Facebook pumping up the volume on what AI can mean to the future of music? You can decide after having a look at what Facebook AI Research scientists have been up to. A number of sites including *The Next Web* have reported that they unveiled a neural network capable of translating music from one style, genre, and set of instruments to another.

You can check out their paper, "A Universal Music Translation Network" by authors Noam Mor, Lior Wolf, Adam Polyak, Yaniv Taigman, Facebook AI Research.

The paper is on arXiv.

Popular Science, another of the sites showing interest, carried a headline suggesting you might take your bad [whistling](#) and make it sound like Mozart.

A video showing the authors' supplementary audio samples lets you hear what they did with samples ranging from symphony, string quartet, to sounds of Africa, Elvis and Rihanna samples and even human whistling.

Tristan Greene, *The Next Web*: "The AI takes one input, such as a symphony orchestra playing Bach, and [translates](#) it into something else, like the same song played on a piano in the style of Beethoven, for example."

In one example, they said they converted the audio of a Mozart symphony performed by an orchestra to an audio in the style of a pianist playing Beethoven.

Basically, a neural network has been put to work to change the style of [music](#). Listening to samples, one wonders what the AI process is like in figuring out how to carry the music from one work to another? Does it involve matched pitch? Memorizing musical notes? Greene said no, their approach is an "unsupervised learning method" using "high-level semantics interpretation." Greene added that you could say "it plays be ear." The method is unsupervised, in that it does not rely on supervision in the form of matched samples between domains or musical transcriptions, said the team.

Call it "strategic confusion." The authors blogged about this on the Facebook Research site. To allow the system to transform music in an unsupervised— improvisational—way, they "intentionally distorted the musical input, with something called a domain confusion network. This prevents the AI from encoding domain-specific information. In other words, the system is forced to ignore the unique aspects of a recorded song's style, genre and instruments, and create translations based the core structure of the music."

Popular Science translated their method, saying "The AI isn't reading musical notes—it's just turning a given audio file into code and then decoding it into a new version."

Greene also translated, explaining that this was "a complex method of auto-encoding that allows the network to process audio from inputs it's never been trained on." *Popular Science* quoted Lior Wolf, one of the co-authors of a new study. "We want to mimic the human ability to hear music, and repeat it, either by whistling or playing an instrument."

In a bigger picture, one can mark the AI attempt to translate styles and instruments as another sure sign of an intersection being crossed between AI and music that can change our pejorative view of "machine" music as inferior and canned.

Music Ally carried an interesting report earlier this month and one of the people noted in the article was author, musician and educator, Marcus O'Dair from Middlesex University. O'Dair named three categories of this technology: AIs composing music; AIs working with humans to co-compose; and AIs that can remix or adapt music for a specific purpose.

He added that the category of AIs composing music was far from science fiction. "I've been in the room where people played two tracks: one's from an AI and one's from a human. And people couldn't tell the

difference," said O'Dair. "I don't think that means an AI is going to write Beethoven's Fifth tomorrow, but for cases like library music… that is certainly interesting."

What's next for Facebook on this AI research into music? The authors of the paper addressed the question in the blog—namely, no plans for a specific product or feature based on this work. Nonetheless, they said that the research was a "strong indicator of how AI could soon power human creativity. From composing whole symphonies with your voice to transforming a simple guitar lick or MIDI tune into layered vocals, this approach could democratize songwriting, and make music production more accessible."

 More information: — research.fb.com/facebook-resea … usical-translations/

— A Universal Music Translation Network, arXiv:1805.07848 [cs.SD] arxiv.org/abs/1805.07848

Abstract
We present a method for translating music across musical instruments, genres, and styles. This method is based on a multi-domain wavenet autoencoder, with a shared encoder and a disentangled latent space that is trained end-to-end on waveforms. Employing a diverse training dataset and large net capacity, the domain-independent encoder allows us to translate even from musical domains that were not seen during training. The method is unsupervised and does not rely on supervision in the form of matched samples between domains or musical transcriptions. We evaluate our method on NSynth, as well as on a dataset collected from professional musicians, and achieve convincing translations, even when translating from whistling, potentially enabling the creation of instrumental music by untrained humans.

© 2018 Tech Xplore