

What makes Bach sound like Bach? New dataset teaches algorithms classical music

December 1 2016, by Jennifer Langston



MusicNet is a new publicly available dataset from UW researchers that labels each note of 330 classical compositions in ways that can teach machine learning algorithms about the basic structure of music. Credit: Yngve Bakken Nilsen, flickr

The composer Johann Sebastian Bach left behind an [incomplete fugue](#) upon his death, either as an unfinished work or perhaps as a puzzle for future composers to solve.

A classical [music](#) dataset released Wednesday by University of Washington researchers—which enables machine learning algorithms to learn the features of classical music from scratch—raises the likelihood that a computer could expertly finish the job.

[MusicNet](#) is the first publicly available large-scale classical music dataset with curated fine-level annotations. It's designed to allow machine learning researchers and algorithms to tackle a wide range of open challenges—from note prediction to automated music transcription to offering listening recommendations based on the structure of a song a person likes, instead of relying on generic tags or what other customers have purchased.

"At a high level, we're interested in what makes music appealing to the ears, how we can better understand composition, or the essence of what makes Bach sound like Bach. It can also help enable practical applications that remain challenging, like automatic transcription of a live performance into a written score," said Sham Kakade, a UW associate professor of computer science and engineering and of statistics.

"We hope MusicNet can spur creativity and practical advances in the fields of machine learning and music composition in many ways," he said.

Described in a paper published Nov. 30 in the arXiv pre-print repository, MusicNet is a collection of 330 freely licensed classical music recordings with annotated labels that indicate the exact start and stop time of each individual note, what instrument plays the note and its position in the composition's metrical structure. It includes more than 1 million individual labels from 34 hours of chamber music performances that can train computer algorithms to deconstruct, understand, predict and reassemble components of classical music.

"The music research community has been working for decades on hand-crafting sophisticated audio features for music analysis. We built MusicNet to give researchers a large labelled dataset to automatically learn more expressive audio features, which show potential to radically change the state-of-the-art for a wide range of music analysis tasks," said Zaid Harchaoui, a UW assistant professor of statistics.

It's similar in design to ImageNet, a public dataset that revolutionized the field of computer vision by labeling basic objects—from penguins to parked cars to people—in millions of photographs. This vast repository of visual data that computer algorithms can learn from has enabled huge strides in everything from image searching to self-driving cars to algorithms that recognize your face in a photo album.

"An enormous amount of the excitement around artificial intelligence in the last five years has been driven by supervised learning with really big datasets, but it hasn't been obvious how to label music," said lead author John Thickstun, a UW computer science and engineering doctoral student.

"You need to be able to say from 3 seconds and 50 milliseconds to 78 milliseconds, this instrument is playing an A. But that's impractical or impossible for even an expert musician to track with that degree of accuracy."

The UW research team overcame that challenge by applying a technique called dynamic time warping—which aligns similar content happening at different speeds—to [classical music](#) performances. This allowed them to synch a real performance, such as Beethoven's 'Serioso' string quartet, to a synthesized version of the same piece that already contained the desired musical notations and scoring in digital form.

Time warping and mapping that digital scoring back onto the original

performance yields the precise timing and details of individual notes that make it easier for machine learning algorithms to learn from musical data.

In their arXiv paper, the UW research team tested the ability of some common end-to-end [deep learning algorithms](#) used in speech recognition and other applications to predict missing notes from compositions. They are making the dataset publicly available so [machine learning](#) researchers and music hobbyists can adapt or develop their own algorithms to advance music transcription, composition, research or recommendations.

"No one's really been able to extract the properties of music in this way, which opens so many opportunities for creative play," said Kakade.

For instance, one could imagine asking your computer to make up a performance that's similar to songs you've listened to, or to hum a melody and tell it to make a fugue on command.

"I'm really interested in the artistic opportunities. Any composer who crafts their art with the assistance of a computer—which includes many modern musicians—could use these tools," said Thickstun. "If the machine has a higher understanding of what they're trying to do, that just gives the artist more power."

More information: Learning Features of Music from Scratch:
arxiv.org/abs/1611.09827

Provided by University of Washington

Citation: What makes Bach sound like Bach? New dataset teaches algorithms classical music (2016, December 1) retrieved 16 April 2024 from <https://techxplore.com/news/2016-12-bach->

<dataset-algorithms-classical-music.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.