Google leverages WaveNet model's gains, sounds seem more natural

October 8 2017, by Nancy Owano

(Tech Xplore)—DeepMind's artificial intelligence talents have been working up capabilities for a consumer product. Sam Shead, Senior Technology Reporter for *Business Insider UK*, said Google applied software developed by DeepMind for use in its virtual assistant.

DeepMind, the AI company, has a version of a WaveNet system for American English and Japanese, according to a blog post published on Wednesday. They said, "we are proud to announce that an updated version of WaveNet is being used to generate the Google Assistant voices for US English and Japanese across all platforms."

"Google has been slow to integrate DeepMind's technology into its products, with just one data centre efficiency project announced so far,

albeit on a global scale," said Shead. "Now the company's WaveNet neural network is being used to generate the Google Assistant voices for US English and Japanese."

Google Assistant is a virtual personal assistant developed by Google.

Pocket-lint described Google Assistant as a voice-controlled smart assistant. "It's considered an upgrade or an extension of Google Now - designed to be personal - while expanding on Google's existing 'OK Google' voice controls."

The DeepMind blog post was from Aäron van den Oord, research scientist, Tom Walters, research scientist, and Trevor Strohman, Google Speech software engineer.

The update they talk about is by the DeepMind WaveNet research and engineering teams, together with the Google Text-to-Speech team.

WaveNet has come a long way in a short time.

Just over a year ago, WaveNet was presented, a deep neural network generating raw audio waveforms and capable of producing speech.

How they built it: A convolutional neural network was trained on a large dataset of speech samples. The goal was more natural-sounding speech than in existing techniques. In their original paper, they said it "creates individual waveforms from scratch, one sample at a time, with 16,000 samples per second and seamless transitions between individual sounds."

As the blog authors put it, "WaveNet showed promise but was not something we could deploy in the real world." It was "too computationally intensive" for use in consumer products. The team got busy to improve the model. They said it now can run "at scale and is the

first product to launch on Google's latest TPU cloud infrastructure."

Key gains:

"The new, improved WaveNet model still generates a raw waveform but at speeds 1,000 times faster than the original model, meaning it requires just 50 milliseconds to create one second of speech."

Ryan Whitwam in *ExtremeTech*: "DeepMind promises a full paper soon that will detail how this was accomplished."

Also, the results are more natural sounding according to tests with human listeners, they blogged.

Whitwam remarked on Friday: "The voice model used in Assistant at launch wasn't bad, but Google just rolled a [vastly](link) improved version of the voices for English and Japanese."

The blog has some interesting summaries of how far the technology has come.

As for current text to speech systems they noted that concatenative TTS not only results in unnatural sounding voices but such systems are hard to modify: a new database needs to be recorded each time there is a shift, such as new emotions or intonations.

To overcome some of these problems, they said an alternative model, parametric TTS, is sometimes used. This approach uses rules and parameters about mouth movements and grammar to deliver—with voices that do not sound altogether natural.

There there's WaveNet.

So, DeepMind, what's next? They said this is just the start for WaveNet. They said they were excited over possibilities that "the power of a voice interface could now unlock for all the world's languages."

  More information: deepmind.com/blog/wavenet-laun … es-google-assistant/

Citation: Google leverages WaveNet model's gains, sounds seem more natural (2017, October 8) retrieved 30 April 2024 from
https://techxplore.com/news/2017-10-google-leverages-wavenet-gains-natural.html