

Synteзатор Lyrebird potrafi naśladować każdy głos

5 maja 2017

Nawet najbardziej naturalne głosy z komputera brzmią sztucznie. Jednak dzięki pracy kanadyjskiej firmy Lyrebird już wkrótce może się to zmienić. Tworzy ona bowiem system sztucznej inteligencji, która jest w stanie naśladować głos dowolnej osoby.

Zaprezentowany w ubiegłym tygodniu system jest w stanie wygenerować tysiące zdań w ciągu sekundy. Działa więc znacznie szybciej niż inne tego typu programy. Potrafi też naśladować dowolny głos, co budzi poważne zastrzeżenia natury etycznej. Wygenerowanie naturalnie brzmiącego głosu to od dawna cel systemów przekładających tekst pisany na mowę. Systemy takie rozwijają Microsoft, Google, Amazon czy Apple. Działają one dzięki łączeniu wyrazów i zdań z wcześniej nagranych pliku z konkretnym głosem. Zmiana głosu wymaga dostępu do pliku z nagraniem każdym możliwym wyrazem potrzebnym do komunikowania się z urządzeniem.

System Lyrebird działa inaczej. Jest on w stanie, na podstawie wysłuchania wielu godzin nagrania audio, nauczyć się wymowy liter, fonemów i wyrazów. Po takim treningu potrafi samodzielnie tworzyć nowe zdania, dodawać intonację czy emocje. Lyrebird korzysta z sieci neuronowych, które działają podobnie do ludzkiego mózgu. Wykorzystują techniki głębokiego uczenia się do przekładania poszczególnych dźwięków na mowę.

Po tym, jak system nauczy się generować mowę, wystarczy mu jednogodzinowe nagranie audio danej osoby, by mówić jej głosem. „Po nauczaniu się wielu głosów dodanie każdego kolejnego odbywa się bardzo szybko. Dlatego też nie potrzebujemy zbyt wielu danych, by nauczyć Lyrebird całkowicie nowego głosu. Oczywiście im więcej danych, tym lepiej, ale już jednogodzinowe

nagranie wystarczy, by system wyłapał wiele cech charakterystycznych nowego głosu” – mówi współtwórca Lyrebird Alexandre de Brebisson, doktorant na Uniwersytecie w Montrealu.

Lyrebird zaprezentowało swój nowy system, prezentując wygenerowane przezeń głosy Baracka Obamy, Donalda Trumpa i Hillary Clinton, które to głosy informują o powstaniu systemu Lyrebird. Trzeba przyznać, że Kanadyjczykom udało się osiągnąć świetne wyniki. Ich system nie jest jednak doskonały. Podczas generowania głosów tworzony jest szum w tle, pozostaje wrażenie pewnej sztuczności, nie zaimplementowano też dźwięków wydawanych przez ruch ust, policzków czy wydychane powietrze.

Najpewniej musimy jeszcze poczekać kilka lat zanim powstaną systemy potrafiące w czasie rzeczywistym naśladować w sposób przekonujący głos dowolnej osoby. Jednak system Lyrebird jest już na tyle doskonały, że wiele osób da się oszukać, jeśli z zaskoczenia usłyszą wygenerowaną przezeń wypowiedź. To oczywiście budzi wiele zastrzeżeń natury etycznej. Doskonałe systemy tego typu będą bowiem w stanie włożyć w usta dowolnej osoby słowa, których nigdy nie wypowiedziała.

Autorstwo: Mariusz Błoński

Na podstawie: ScientificAmerican.com

Źródło: KopalniaWiedzy.pl