

AI Outperforms Humans in Speech Recognition

KIT Researchers Have Developed the First Speech Recognition System Worldwide that Works Better than Humans and Is Quicker than Other AIs



Thanks to its superior speech recognition system, KIT's Lecture Translator will provide better results with minimum latency in future. (Photo: KIT)

Following a conversation and transcribing it precisely is one of the biggest challenges in artificial intelligence (AI) research. For the first time now, researchers of Karlsruhe Institute of Technology (KIT) have succeeded in developing a computer system that outperforms humans in recognizing such spontaneously spoken language with minimum latency. This is reported on the Internet platform [ArXiv.org](https://arxiv.org).

“When people talk to each other, there are stops, stutterings, hesitations, such as “er” or “hmmm”, laughs and coughs,” says Alex Waibel, Professor for Informatics at KIT. “Often, words are pronounced unclearly.” This makes it difficult even for people to make accurate notes of a conversation. “And so far, this has been even more difficult for AI,” adds the speech recognition expert. KIT scientists and staff of KITES, a start-up company from KIT, have now programmed a computer system that executes this task better than humans and quicker than other systems.

Waibel already developed an automatic live translator that directly translates university lectures from German or English into the languages spoken by foreign students. This “Lecture Translator” has



KIT Information · Systems · Technologies Center

Monika Landgraf
Chief Press Officer,
Head of Corp. Communications

Kaiserstraße 12
76131 Karlsruhe
Phone: +49 721 608-41105
Email: presse@kit.edu

Press contact:

Dr. Felix Mescoli
Press Officer
Phone: +49 721 608-41171
Email: felix.mescoli@kit.edu

Further material:

Link to the paper:
<https://arxiv.org/abs/2010.03449>

been used in the lecture halls of KIT since 2012. "Recognition of spontaneous speech is the most important component of this system," Waibel explains, "as errors and delays in recognition make the translation incomprehensible. On conversational speech, the human error rate amounts to about 5.5%. Our system now reaches 5.0%." Apart from precision, however, the speed of the system to produce output is just as important so students can follow the lecture live. The researchers have now succeeded in reducing this latency to one second. This is the smallest reported latency reached by a speech recognition system of this quality to date, says Waibel.

Error rate and latency are measured using the standardized and internationally recognized, scientific "switchboard-benchmark" test. This benchmark (defined by US NIST) is widely used by international AI researchers in their competition to build a machine that comes close to humans in recognizing spontaneous speech under comparable conditions, or even outperforming them.

According to Waibel, fast, high accuracy speech recognition is an essential step for further downstream processing. It enables dialog, translation, and other AI modules to provide better voice based interaction with machines.

More about the KIT Information · Systems · Technologies Center:
<http://www.kcist.kit.edu>

Being "The Research University in the Helmholtz-Association," KIT creates and imparts knowledge for the society and the environment. It is the objective to make significant contributions to the global challenges in the fields of energy, mobility and information. For this, about 9,300 employees cooperate in a broad range of disciplines in natural sciences, engineering sciences, economics, and the humanities and social sciences. KIT prepares its 24,400 students for responsible tasks in society, industry, and science by offering research-based study programs. Innovation efforts at KIT build a bridge between important scientific findings and their application for the benefit of society, economic prosperity, and the preservation of our natural basis of life. KIT is one of the German universities of excellence.

This press release is available on the internet at www.kit.edu.

The photo in the best quality available to us may be downloaded under www.kit.edu or requested by mail to presse@kit.edu or phone +49 721 608-41105. The photo may be used in the context given above exclusively.