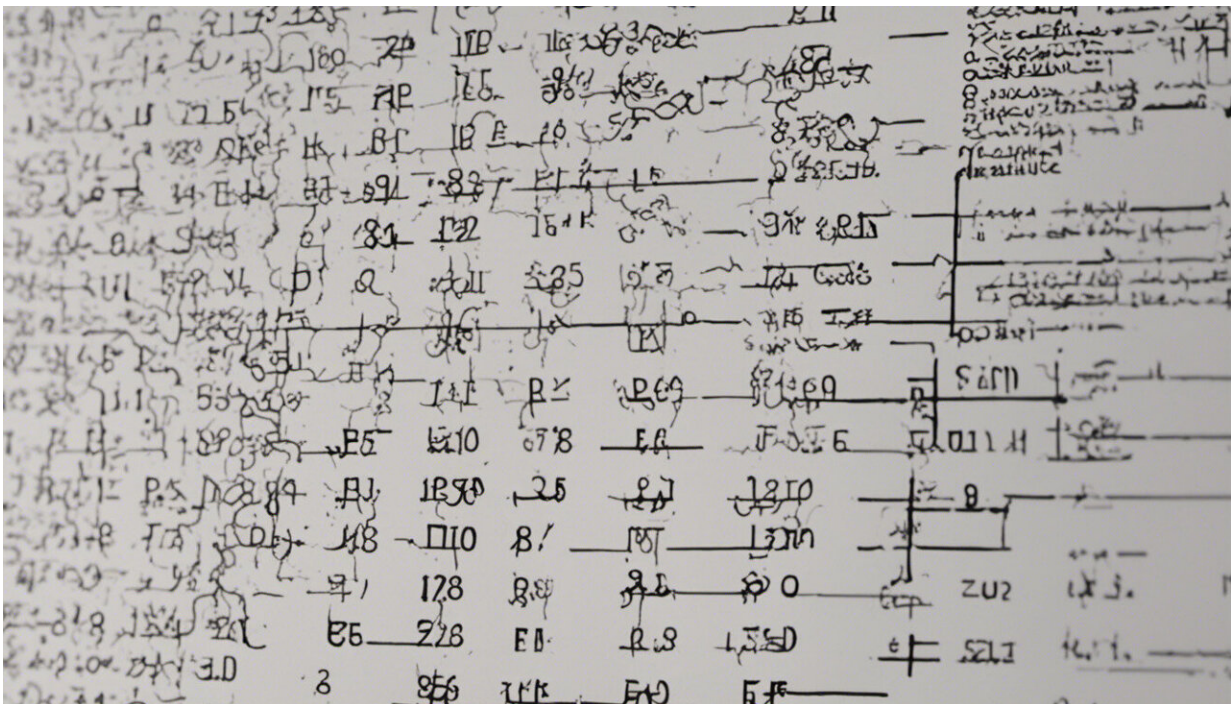


# New computer software program excels at lip reading

March 17 2017



Credit: AI-generated image ([disclaimer](#))

A new computer software program has the potential to lip-read more accurately than people and to help those with hearing loss, Oxford University researchers have found.

Watch, Attend and Spell (WAS), is a new artificial intelligence (AI)

software system that has been developed by Oxford, in collaboration with the company DeepMind.

The AI system uses computer vision and machine learning methods to learn how to lip-read from a dataset made up of more than 5,000 hours of TV footage, gathered from six different programmes including Newsnight, BBC Breakfast and Question Time. The videos contained more than 118,000 sentences in total, and a vocabulary of 17,500 words.

The research team compared the ability of the machine and a human expert to work out what was being said in the silent video by focusing solely on each speaker's lip movements. They found that the software system was more accurate compared to the professional. The human lip-reader correctly read 12 per cent of words, while the WAS software recognised 50 per cent of the words in the dataset, without error. The machine's mistakes were small, including things like missing an "s" at the end of a word, or single letter misspellings.

The [software](#) could support a number of developments, including helping the hard of hearing to navigate the world around them. Speaking on the tech's core value, Jesal Vishnuram, Action on Hearing Loss Technology Research Manager, said: 'Action on Hearing Loss welcomes the development of new technology that helps people who are deaf or have a [hearing loss](#) to have better access to television through superior real-time subtitling.

'It is great to see research being conducted in this area, with new breakthroughs welcomed by Action on Hearing Loss by improving accessibility for people with a hearing loss. AI lip-reading technology would be able to enhance the accuracy and speed of speech-to-text especially in noisy environments and we encourage further research in this area and look forward to seeing new advances being made.'

Commenting on the potential uses for WAS Joon Son Chung, lead-author of the study and a graduate student at Oxford's Department of Engineering, said: 'Lip-reading is an impressive and challenging skill, so WAS can hopefully offer support to this task - for example, suggesting hypotheses for professional lip readers to verify using their expertise. There are also a host of other applications, such as dictating instructions to a phone in a noisy environment, dubbing archival silent films, resolving multi-talker simultaneous speech and improving the performance of automated speech recognition in general.'

The research team comprised of Joon Son Chung and Professor Andrew Zisserman at Oxford, where the research was carried out, together with Dr Andrew Senior and Dr Oriol Vinyals at DeepMind. Professor Zisserman commented 'this project really benefitted by being able to bring together the expertise from Oxford and DeepMind'.

Provided by University of Oxford

Citation: New computer software program excels at lip reading (2017, March 17) retrieved 4 May 2024 from <https://phys.org/news/2017-03-software-excels-lip.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.