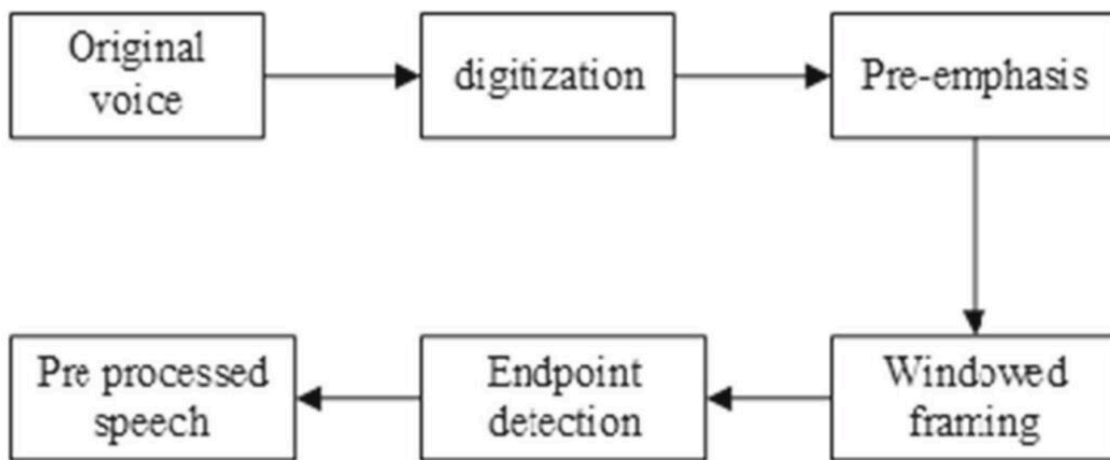# Using deep learning to detect depression from speech

February 7 2023, by Ingrid Fadelli

Flowchart of the model's front-end processing of voice signals. Credit: Tian, Zhu and Jing

Artificial intelligence (AI) tools have achieved promising results on numerous tasks and could soon assist professionals in various settings. In recent years, computer scientists have been exploring the potential of these tools for detecting signs of different physical and psychiatric conditions.

Depression is one of the most widespread psychiatric disorders,

affecting approximately 9.5% of American adults every year. Tools that can automatically detect signs of depression might help to reduce suicide rates, as they would allow doctors to promptly identify people in need of psychological support.

Researchers at Jinhua Advanced Research Institute and Harbin University of Science and Technology have recently developed a deep learning algorithm that could detect depression from a person's speech. This model, introduced in a paper published in *Mobile Networks and Applications*, was trained to recognize emotions in human speech by analyzing different relevant features.

"A multi-information joint decision algorithm model is established by means of emotion recognition," Han Tian, Zhang Zhu and Xu Jing wrote in their paper. "The model is used to analyze the representative data of the subjects, and to assist in diagnosis of whether the subjects have depression."

Tian and his colleagues trained their model on the DAIC-WOZ dataset, a collection of audio and 3D face expressions of patients diagnosed with depressive disorder and of people without depression. These audio recordings and facial expressions were collected during interviews led by a virtual agent, who asked different questions about the interviewee's mood and life.

"On the basis of exploring the speech characteristics of people with depressive disorder, this paper conducts an in-depth study of speech assisted depression diagnosis based on the speech data in the DAIC-WOZ dataset," Tian, Zhu, and Jian wrote in their paper. "First, the speech information is preprocessed, including speech signal pre-emphasis, framing windowing, endpoint detection, noise reduction, etc. Secondly, OpenSmile is used to extract the features of speech signals, and the speech features that the features can reflect are studied and

analyzed in depth."

To extract relevant features from voice recordings, the team's model uses OpenSmile (open-source speech and music interpretation by large-space extraction). This is a toolkit often used by computer scientists to extract features from audio clips and classify these clips.

The researchers used this tool to extract individual speech features and combinations of features that are commonly found in the speech of patients diagnosed with depression. Subsequently, they used a technique known as principal component analysis to reduce the set of features that they extracted.

Tian, Zhu and Jian evaluated their model in a series of tests, where they assessed its ability to detect depressed and non-depressed people from recordings of their voice. Their framework achieved remarkable results, detecting depression with an accuracy of 87% in male patients and 87.5% in female patients.

In the future, the deep learning algorithm developed by this team of researchers could be an additional assistive tool for human psychiatrists and doctors, along with other well-established diagnostic tools. In addition, this study could inspire the development of similar AI tools for detecting signs of psychiatric disorders from speech.

May 2024 from https://medicalxpress.com/news/2023-02-deep-depression-speech.html