AI can detect depression in a child's speech

May 6 2019

A machine learning algorithm can detect signs of anxiety and depression in the speech patterns of young children, potentially providing a fast and easy way of diagnosing conditions that are difficult to spot and often overlooked in young people, according to new research published in the *Journal of Biomedical and Health Informatics*. Credit: Anthony Kelly

A machine learning algorithm can detect signs of anxiety and depression

in the speech patterns of young children, potentially providing a fast and easy way of diagnosing conditions that are difficult to spot and often overlooked in young people, according to new research published in the *Journal of Biomedical and Health Informatics.*

Around one in five children suffer from anxiety and depression, collectively known as "internalizing disorders." But because children under the age of eight can't reliably articulate their emotional suffering, adults need to be able to infer their mental state, and recognise potential mental health problems. Waiting lists for appointments with psychologists, insurance issues, and failure to recognise the symptoms by parents all contribute to children missing out on vital treatment.

"We need quick, objective tests to catch kids when they are suffering," says Ellen McGinnis, a clinical psychologist at the University of Vermont Medical Center's Vermont Center for Children, Youth and Families and lead author of the study. "The majority of kids under eight are undiagnosed."

Early diagnosis is critical because children respond well to treatment while their brains are still developing, but if they are left untreated they are at greater risk of substance abuse and suicide later in life. Standard diagnosis involves a 60-90 minute semi-structured interview with a trained clinician and their primary care-giver. McGinnis, along with University of Vermont biomedical engineer and study senior author Ryan McGinnis, has been looking for ways to use artificial intelligence and machine learning to make diagnosis faster and more reliable.

The researchers used an adapted version of a mood induction task called the Trier-Social Stress Task, which is intended to cause feelings of stress and anxiety in the subject. A group of 71 children between the ages of three and eight were asked to improvise a three-minute story, and told that they would be judged based on how interesting it was. The

researcher acting as the judge remained stern throughout the speech, and gave only neutral or negative feedback. After 90 seconds, and again with 30 seconds left, a buzzer would sound and the judge would tell them how much time was left.

"The task is designed to be stressful, and to put them in the mindset that someone was judging them," says Ellen McGinnis.

The children were also diagnosed using a structured clinical interview and parent questionnaire, both well-established ways of identifying internalizing disorders in children.

The researchers used a machine learning algorithm to analyze statistical features of the audio recordings of each kid's story and relate them to the child's diagnosis. They found the algorithm was highly successful at diagnosing children, and that the middle phase of the recordings, between the two buzzers, was the most predictive of a diagnosis.

"The algorithm was able to identify children with a diagnosis of an internalizing disorder with 80% accuracy, and in most cases that compared really well to the accuracy of the parent checklist," says Ryan McGinnis. It can also give the results much more quickly—the algorithm requires just a few seconds of processing time once the task is complete to provide a diagnosis.

The algorithm identified eight different audio features of the children's speech, but three in particular stood out as highly indicative of internalizing disorders: low-pitched voices, with repeatable speech inflections and content, and a higher-pitched response to the surprising buzzer. Ellen McGinnis says these features fit well with what you might expect from someone suffering from depression. "A low-pitched voice and repeatable speech elements mirrors what we think about when we think about depression: speaking in a monotone voice, repeating what

you're saying," says Ellen McGinnis.

The higher-pitched response to the buzzer is also similar to the response the researchers found in their previous work, where children with internalizing disorders were found to exhibit a larger turning-away response from a fearful stimulus in a fear induction task.

The voice analysis has a similar accuracy in diagnosis to the motion analysis in that earlier work, but Ryan McGinnis thinks it would be much easier to use in a clinical setting. The fear task requires a darkened room, toy snake, motion sensors attached to the child and a guide, while the voice task only needs a judge, a way to record speech and a buzzer to interrupt. "This would be more feasible to deploy," he says.

Ellen McGinnis says the next step will be to develop the speech analysis algorithm into a universal screening tool for clinical use, perhaps via a smartphone app that could record and analyze results immediately. The voice analysis could also be combined with the motion analysis into a battery of technology-assisted diagnostic tools, to help identify children at risk of anxiety and depression before even their parents suspect that anything is wrong.

Provided by University of Vermont